# ACT® TECHNICAL MANUAL

SEPTEMBER 2026

## Commitment to Fair Testing

ACT endorses and is committed to complying with the *Standards for Educational and Psychological Testing* (American Educational Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in Education [NCME], 2014). ACT also endorses the *Code of Fair Testing Practices in Education* (Joint Committee on Testing Practices, 2004), which discusses the obligations of those who develop, administer, or use educational tests and test data, particularly in the following four areas: developing and selecting appropriate tests, administering and scoring tests, reporting and interpreting test results, and communicating with test takers. ACT also endorses and is committed to complying with the *Code of Professional Responsibilities in Educational Measurement* (NCME Ad Hoc Committee on the Development of a Code of Ethics, 2023), which specifies the professional responsibilities of those involved with various aspects of tests, including their development, marketing, interpretation, and use.

# Table of Contents

Commitment to Fair Testing ..... i
Table of Contents ..... ii
List of Tables ..... iv
List of Figures ..... vi
Preface ..... viii
Chapter 1: The ACT® ..... 1
ACT's Mission ..... 1
1.1 Philosophical Basis for the ACT ..... 1
1.2 Overview of the ACT ..... 1
1.3 Purposes, Claims, Interpretations, and Uses of the ACT ..... 2
1.4 Evidence-Based Design of the ACT Test ..... 5
1.5 ACT's Commitment to Fair Testing ..... 7
1.6 The Population Served by the ACT ..... 8
1.7 Test Preparation ..... 9
Chapter 2: The ACT Test Development ..... 11
2.1 Overview ..... 11
2.2 Description of the ACT Tests ..... 12
2.3 The ACT National Curriculum Survey ..... 12
2.4 Test Development Procedures ..... 14
2.5 Test Development Procedures for the Writing Test ..... 22
2.6 ACT Scores ..... 23
Chapter 3: Content Specifications ..... 26
3.1 Overview ..... 26
3.2 English Test ..... 26
3.3 Mathematics Test ..... 28
3.4 Reading Test ..... 32
3.5 Science Test ..... 34
3.6 Writing Test ..... 37
Chapter 4: Test Administration, Test Security, and Accessibility and Accommodations ..... 40
4.1 Test Administration Overview ..... 40
4.2 Test Security ..... 41
4.3 Information Security ..... 42
4.4 Test Administration and Accessibility Levels of Support ..... 43
Chapter 5: Scoring and Reporting ..... 47
5.1 Overview ..... 47
5.2 Test Section Scores, Composite Scores, STEM and ELA Scores, and Superscores ..... 47
5.3 Detailed Performance Description ..... 79
5.4 Progress Toward the ACT WorkKeys National Career Readiness Certificate Indicator ..... 81
5.5 ACT College and Career Readiness Standards ..... 83
5.6 ACT College Readiness Benchmarks ..... 87

Chapter 6: Scaling, Equating, and Technical Characteristics ....................................................90
6.1 Scaling, Linking, and Equating the ACT English, Mathematics, Reading, and Science Test Sections ....................................................90
6.2 Scaling and Equating the ACT Writing Test Section to Calculate ELA Scores ....................94

Chapter 7: Validity Evidence for the ACT Tests....................................................105
7.1 Using ACT Scores to Measure Educational Achievement....................................................105
7.2 Using ACT Scores to Support College Admissions Decisions ....................................................128
7.3 Using ACT Scores to Support Course Placement Decisions ....................................................145
7.4 Using ACT Scores to Evaluate Students' Likelihood of College Success in the First Year and Beyond ....................................................156
7.5 Using ACT Scores to Assist With Program Evaluation ....................................................170

Chapter 8: Growth Models Using ACT Test Scores ....................................................174
8.1 Distinguishing Gain-Based Models From Conditional Status Models....................................................174
8.2 ACT Growth Modeling Resources ....................................................175
8.3 Explaining Variation in Student Growth ....................................................177
8.4 Using Growth Models for Evaluating Programs and School Effectiveness ....................181
8.5 ACT Retest Statistics....................................................182

Chapter 9: Other ACT Components ....................................................184
9.1 The ACT Interest Inventory....................................................184
9.2 The MyACT Profile ....................................................188

# List of Tables


Table 1.1. Demographic Characteristics of National ACT Testers, June 2025 to April 2026 ....9

Table 2.1. ACT National Curriculum Survey 2020 Respondents ....14

Table 2.2. Difficulty[a] Distributions and Mean Discrimination[b] Indices for ACT Test Items, June 2025 to April 2026 ....19

Table 2.3. ACT Test Items Exhibiting Differential Item Functioning, June 2025 to April 2026 ....21

Table 3.1. DOK Level Descriptions for English ....27

Table 3.2. Specification Ranges by Reporting Category for English (Scored Items Only) ....28

Table 3.3. DOK Level Descriptions for Mathematics ....29

Table 3.4. Specification Ranges by Reporting Category for Mathematics (Scored Items Only) ....32

Table 3.5. DOK Level Descriptions for Reading ....33

Table 3.6. Specification Ranges by Reporting Category for Reading (Scored Items Only) ....34

Table 3.7. DOK Level Descriptions for Science ....35

Table 3.8. Specification Ranges by Reporting Category for Science (Scored Items Only) ....36

Table 3.9. Specification Ranges by Science Content Area (Scored Items Only) ....37

Table 3.10. DOK Level Description for Writing ....37

Table 5.1. Writing Test Analytic Scoring Rubric ....51

Table 5.2. Sample of Quality Reports ....73

Table 5.3. Summary Statistics of the ACT Test Score Distributions ....77

Table 5.4. Summary Statistics of the Best ACT Test Section Score and ACT Superscore Distributions ....77

Table 5.5. Scale Score Covariance Matrix for Multiple-Choice Tests From One ACT Test Form ....78

Table 5.6. Range of Effective Weights for Composite, STEM, and ELA Scores ....78

Table 5.7. Correlations Among the ACT Test Scores ....78

Table 5.8. Summary Statistics of the ACT Writing and Writing Domain Score Distributions ....81

Table 5.9. Correlations Among the ACT Writing and Writing Domain Scores ....81

Table 5.10. ACT College Readiness Benchmarks ....87

Table 6.1. Summary Statistics of Scale Score Reliability and SEMs for the ACT Test Scores ....96

Table 6.2. Summary Statistics of Raw Score Reliability and SEM for the ACT Reporting Categories ....97

Table 6.3. Agreement Rates for the ACT Writing Domain Scores ..........94
Table 6.4. Classification Consistency for the ACT College Readiness Benchmarks ..........97
Table 6.6. Composite Score Ranges for the WorkKeys NCRC Levels ..........99

Table 7.1. Use of Targeted Skills Across Domains ..........109
Table 7.2. Student Ability to Understand What the Question Was Asking by Test Section ..........109
Table 7.3. Students' Rating of Their Ability to Show What They Know and Can Do by Test Section ..........110
Table 7.4. College Readiness Benchmark Attainment Percentages and Average ACT Scores by Common Course Patterns, ACT-Tested High School Graduating Class of 2021 ..........111
Table 7.5. Average ACT Scores by Academic Preparation, 2012–2016 ..........112
Table 7.6. Average ACT Score by HSGPA Ranges, 2015–2016 ..........115
Table 7.7. Correlations Among ACT Scores and Background Characteristics ..........117
Table 7.8. Mean HSGPA and Mean ACT Composite Score, ACT-Tested Public School Graduates, 2010–2021 ..........125
Table 7.9. Change in Mean Subject-Area GPA and Mean ACT Subject Score Between 2010 and 2022, ACT-Tested Public School Graduates ..........125
Table 7.10. Meta-Analysis of Multi-Institutional Data—Correlations With FYGPA, Overall Analyses ..........130
Table 7.11. General Dominance Statistics for Predictors of FYGPA ..........132
Table 7.12. Median Statistics for Predicting 3.0 FGYPA or Higher by Race/Ethnicity Across 236 Institutions ..........141
Table 7.13. Average ACT Scores for Accommodated and Non-Accommodated Examinees, by Disability Category ..........143
Table 7.14. Decision-Based Validity Statistics for Course Placement Using ACT Scores (Success Criterion = B or Higher Grade) ..........149
Table 7.15. Median Placement Statistics for ACT Scores and HSGPA as Predictors at Community Colleges ..........151
Table 7.16. Probability of Success in Mathematics 100 Given ACT Mathematics Score ..........153
Table 7.17. Decision-Based Statistics for Placement Based on ACT Mathematics Score ..........155
Table 7.18. Percent Chance of First-Year College Outcomes by Number of ACT College Readiness Benchmarks Met ..........161
Table 7.19. ACT/CAAP Test Score Correlations ..........162
Table 7.20. Average ACT CAAP Test Score by ACT Benchmark Attainment ..........162

# List of Figures

Figure 1.1. The Full Picture: Evidence and Validity ....... 7

Figure 5.1. Overall Score and Percentile Rank on a Sample Interactive Score Report on MyACT ....... 48

Figure 5.2. Overall Score and Percentile Rank on a Sample ACT High School Score Report ....... 48

Figure 5.3. Detailed Results on a Sample Interactive Score Report on MyACT ....... 79

Figure 5.4. Detailed Results on a Sample ACT High School Score Report ....... 80

Figure 5.5. ACT High School Report Indicator of Progress Toward the ACT WorkKeys National Career Readiness Certificate ....... 82

Figure 6.1. CSEM for Multiple-Choice Test Scores ....... 98

Figure 7.1. Unadjusted and Adjusted Mean Differences in ACT Scores by Family Income ....... 119

Figure 7.2. ACT Composite Cumulative Percentages for 2016 ACT-Tested High School Graduates and Talent Search 6th-, 7th-, and 8th-Grade Students ....... 122

Figure 7.3. Predicted FYGPA by HSGPA at Three Levels of School Grade Inflation, ACT Composite = 21 ....... 126

Figure 7.4. Conditional $R^2$ of HSGPA and ACT Composite Score From 2010 to 2021 ....... 131

Figure 7.5. Probability of C or Higher FYGPA and Accuracy Rate ....... 135

Figure 7.6. Probabilities of Earning a 3.0 or Higher FYGPA and Being Retained Through the First Year Based on HSGPA and ACT Composite Score ....... 138

Figure 7.7. Estimated Probabilities of Achieving Specific FYGPA Levels Based on ACT Composite Score, by Race/Ethnicity ....... 140

Figure 7.8. Estimated Probabilities of Achieving Specific FYGPA Levels Based on HSGPA, by Race/Ethnicity ....... 140

Figure 7.9. Probability of Success in Mathematics 100 Given ACT Mathematics Score ....... 154

Figure 7.10. College Enrollment Rates by ACT College Readiness Benchmark Attainment ....... 158

Figure 7.11. Students' Chances of Earning a B or Higher in First-Year College Courses by ACT College Readiness Benchmark Attainment at a Typical Institution ....... 159

Figure 7.12. Students' Chances of Achieving a 3.0 or Higher FYGPA by ACT College Readiness Benchmark Attainment at a Typical Institution ....... 160

Figure 7.13. Students' Chances of Remaining Enrolled at the Initial Institution in Year 2 by ACT College Readiness Benchmark Attainment ....... 160

Figure 7.14. Percentages Earning a Cumulative College GPA Greater Than 3.0 by ACT College Readiness Benchmark Attainment for Students Taking ACT CAAP During Sophomore Year and the ACT in High School ....... 163

Figure 7.15. Path Analysis of the Effects of ACT Composite Score and HSGPA on CGPA With FYGPA as a Mediator ....... 164

Figure 7.16. Probability of Degree Completion Based on ACT Composite Score....................166
Figure 7.17. Probability of Bachelor's Degree Completion Within 6 Years by HSGPA and ACT Composite Score ............................................................................167
Figure 7.18. Path Analysis of Effect of ACT Composite Score and HSGPA on Probability of Degree Completion by Year 4 With FYGPA as a Mediator .................................168

Figure 8.1. Average Gain in ACT Composite Score by Months of Instruction........................178

Figure 9.1. Example ACT Interest Inventory Results From the Student Report on MyACT ...185
Figure 9.2. The ACT Career Map.............................................................................186
Figure 9.3. Example Interest-Major Fit Bar From the ACT College Report............................188

# Preface

The principal purpose of the *ACT Technical Manual* is to document the technical characteristics of the ACT® test according to its intended uses and interpretations. The *ACT Technical Manual* documents the validity evidence that supports appropriate interpretations of test scores, and it describes various content-related and psychometric aspects of the ACT. The manual articulates multiple test design and development processes, documents how ACT builds the test in line with the validity argument, and explains how concepts such as construct validity, fairness, and accessibility are maintained throughout the process. The technical manual also describes ACT's routine analyses designed to support continuous improvement and the research that ACT undertakes to ensure that the program remains both psychometrically and educationally sound.

We encourage those who want more detailed information on a topic discussed in this manual, or on a related topic, to contact ACT. Please direct comments or inquiries to The ACT Test: Contact Us.

# Chapter 1
# The ACT®

## ACT's Mission

ACT has been dedicated to improving college and career readiness for all students since its inception in 1959. ACT's renowned longitudinal system of assessments, with the ACT® test as a capstone, has provided students, educators, and policymakers with unparalleled measures of readiness, advancing our mission of helping people achieve education and workplace success.

## 1.1 Philosophical Basis for the ACT

Underlying the ACT is the belief that students' preparation for college and the workplace is best assessed by measuring, as directly as possible, the skills learned in high school that are required for success in college-level courses. The required academic skills can be assessed most directly by reproducing, as faithfully as possible, the complexity of the work students do in the classroom. Therefore, ACT's tests of educational achievement are designed to determine how skillfully students solve problems, grasp implied meanings, draw inferences, evaluate ideas, and make judgments in subject-matter areas important to success in college.

The ACT is oriented toward the general content areas of college and high school instructional programs. The test questions require students to integrate the knowledge and skills they possess in major curriculum areas with the information provided by the test. Thus, scores on the test are directly related to students' educational progress in curriculum-related areas and possess meaning that is readily grasped by students, parents, and educators.

The constructs measured by the ACT section tests are supported by multiple sources of validity evidence (see Chapter 7). For example, ACT has, for many years, collected longitudinal statistical evidence backing the strong relationship between student performance on the section tests and student performance in entry-level college courses in corresponding subjects. More recent methodologies such as cognitive labs have further confirmed this relationship.

The ACT requires students to apply critical thinking skills when comprehending complex texts, analyzing data displays showing the results of scientific experiments, producing effective argumentative writing, and solving sophisticated mathematics problems. Accordingly, long-term learning in school, rather than short-term cramming and coaching, becomes the best form of test preparation. Because the ACT is an educational achievement test that aligns with rigorous high school coursework, an ACT score reflects a level of achievement that has been reached through hard work and study.

## 1.2 Overview of the ACT

The ACT emphasizes students' academic preparedness by directly addressing the content domains students must master to achieve college and career readiness. The main component of the ACT is a standardized battery of three tests of educational achievement—English, mathematics, and reading—along with optional science and writing tests. Through ACT's online registration platform (MyACT), ACT also collects information about students' high school courses and grades, educational and career aspirations, extracurricular activities, and educational needs.

The ACT provides information about how well a student performs compared to other students. It also provides standards-based interpretations through the ACT® College and Career Readiness Standards (CCRS)—empirically derived descriptions of the essential skills and knowledge students need in order to become ready for college and career success. Using the CCRS, secondary educators can pinpoint the skills students have and those they are ready to learn next. The CCRS clarify college expectations in terms that high school teachers understand. The CCRS also offer teachers guidance for improving instruction to help correct student deficiencies in specific areas. The ACT® College Readiness Benchmarks are the minimum scores associated with a high likelihood of postsecondary success in each content area. Together, the CCRS and the Benchmarks provide specific insights into student readiness for college and career. Chapter 5 gives details about the CCRS and Benchmarks.

This edition of the *ACT Technical Manual* reflects the 2025 ACT enhancements, including optional science testing, reduced test length, and embedded field-test items. These enhancements are reflected in updated test specifications and a redefined Composite score based on the English, mathematics, and reading scores. References throughout this manual to legacy ACT forms or scores refer to versions of the ACT that were administered prior to these enhancements.

## 1.3 Purposes, Claims, Interpretations, and Uses of the ACT

The purposes, claims, interpretations, and uses of the ACT are reflected in a theory of action that integrates evidence supporting content validity (academic research, curriculum information, and academic standards) with evidence supporting predictive validity (empirical data). The theory of action begins by answering fundamental questions about the purpose, users, uses, benefits, claims, interpretations, and outcomes of the test.

***Intended Purpose.*** The primary purpose of the ACT is to measure students' level of college and career readiness in core academic areas. ACT testing is intended to help high school students develop postsecondary educational plans and to help postsecondary educational institutions meet the needs of their students.

In service of the intended purpose, the ACT provides an overall Composite score and scores for each of the section tests and the optional science and writing tests. The test also provides a measure of students' STEM skills (by combining mathematics and science scores if both tests were taken) and an ELA score (by combining English, reading, and writing scores if all three tests were taken). The ACT also provides information about student achievement at a more detailed level through the reporting category scores on each test section.

***Intended Users.*** Primary intended users of the ACT test include high school students (typically in Grades 11 and 12), the educational agencies or organizations supporting the academic preparation of these students (i.e., schools, districts, and states), postsecondary institutions, and talent recognition and scholarship agencies.

***Intended Uses.*** ACT test data, test scores, and score interpretations have several intended uses. Students use their results to plan for further education and explore careers based on their skills, interests, and aspirations. High schools use ACT data in academic advising and counseling, evaluation studies, accreditation documentation, and public relations. Postsecondary institutions use ACT results to support admission and course placement decisions. States use the ACT as part of their statewide assessment systems to measure students' educational achievement, to monitor educational improvement and achievement gaps over time, and to meet federal accountability requirements. Many private, state, and national agencies that provide scholarships, loans, and other types of financial assistance use ACT test scores to help assess students' academic qualifications. Agencies also use ACT data to identify academically talented students as early as middle school.

***Intended Benefits.*** The ACT test benefits its users by

- allowing students to demonstrate the knowledge and skills gained through educational coursework in English, mathematics, reading, science, and writing;
- providing students with a profile of their relative strengths and weaknesses in the section areas assessed by the test, thereby informing students about what they know and can do (based on the College and Career Readiness Standards);
- providing parents with insights about their students' knowledge and skills;
- providing educators (in schools, districts, and states) with information about their students' knowledge and skills;
- encouraging students to better prepare for college and careers by taking courses linked to positive postsecondary outcomes;
- indicating whether a student is likely ready for college-level coursework or a work training program (based on the College Readiness Benchmarks and the Progress Toward the ACT® WorkKeys® National Career Readiness Certificate™ (NCRC®) indicator); and
- providing colleges, talent identification programs, and scholarship agencies with information about students' level of achievement in the areas assessed by the test.

***Interpretations and Claims.*** The interpretations and claims of the ACT include the following:

- The ACT measures academic knowledge and skills that are acquired in high school and are important for college-level coursework in English, mathematics, reading, science, and writing.

- ACT scores can be used in combination with other relevant measures to estimate students' likelihood of success in college during the first year and beyond and to help inform college admission, course placement, and remediation decisions.
- ACT scores can be used in aggregate for monitoring educational improvement and achievement gaps over time, as well as assisting with evaluating the effectiveness of school and district programs when a school administers the ACT to all its students.
- MyACT includes the ACT Interest Inventory (ACT, 2023), which is based on research about career planning, to point students toward a range of good-fit options to consider. In the process of exploration, students can focus on educational and occupational options that are relevant to future satisfaction and success. The ACT Interest Inventory results, when used in conjunction with ACT test scores, provide a more holistic picture of the student's educational development and career-relevant motivations.

***Intended Outcomes.*** Using the results of the ACT in conjunction with other academic and nonacademic measures can help in the following ways:

- Students, parents, and educators can identify academic knowledge and skill areas in which students might benefit from additional instruction and supports while still in high school; in this way, they can better prepare for college and career and avoid taking remedial or developmental courses in their first year of college.
- Students can consider their academic strengths and weaknesses, as well as their interests as measured by the ACT Interest Inventory (ACT, 2023) or through ACT's Recruit Me service, to expand their educational and occupational exploration beyond the options they initially considered.
- Schools and districts can increase student awareness of college opportunities when all students take the ACT through state or district testing.
- Schools and districts can evaluate student growth and identify gaps in educational achievement in order to better understand which school programs are effective in preparing all students for college and career.
- Postsecondary institutions can admit students who are likely to enroll at the institution and, once enrolled, likely to succeed in their courses and complete a degree at the institution.
- Postsecondary institutions can place students in first-year college courses in which they are most likely to succeed.
- Postsecondary institutions can identify students early on who are most likely to struggle academically, who may be at risk of dropping out, and who may benefit from

institutional academic services and supports in order to successfully transition from high school to college.

## 1.4 Evidence-Based Design of the ACT Test

The design of the ACT test emerges from an evidence-based research and data-collection process that ensures that items and test forms elicit the evidence necessary to support the claims of the ACT. For example, content and item specifications and test blueprints influence the technical quality of test items and forms. The ACT design is informed by several factors, including the following:

- subject matter experts (SMEs), including external expert panels and internal content development staff
- academic research on skill targets, sequencing of skills, and grade placement
- data and evidence of student understanding collected from the ACT test
- the ACT® National Curriculum Survey®
- a survey of standards frameworks such as the ACT College and Career Readiness Standards, the Next Generation Science Standards, and other college and career readiness standards
- cognitive lab research including think-alouds, eyetracking, and interviews

The ACT National Curriculum Survey provides empirical validation evidence related to the content of the tests. The most recent survey was released in 2020 and included responses from thousands of educators, from K–12 to college, in ELA, mathematics, and science (ACT, 2020a). New survey data were collected during the 2025–26 school year, and a new National Curriculum Survey Report is scheduled for release during the 2026–27 school year.

The 2020 ACT National Curriculum Survey also includes workforce supervisors and employees, whose responses are reviewed by SMEs and used to identify the most critical skills and knowledge required for career readiness.

The validity argument is further supported with criterion-related longitudinal evidence from students who complete the ACT and then go on to colleges (2-year and 4-year) and career-training programs.

While SMEs can identify copious skills covered by a typical high school curriculum, not all skills and knowledge are essential for postsecondary success, nor will measuring every skill help identify lower- and higher-achieving students. For example, some skills essential for success may be attained by more than 95 percent of students continuing on to postsecondary education; putting items that measure such skills on a test only increases test length without contributing to predicting postsecondary success.

Similarly, ACT research demonstrates that there are often discrepancies between skills high school educators see as relevant to success and the expectations and experience of college faculty. Again, ACT uses data from a national sample of institutions, academic programs, and college majors to prioritize the skills and knowledge clearly linked to student success.

ACT supplements these other sources of data with subject matter expertise. ACT's test development staff has extensive classroom experience in the subjects tested by the ACT.

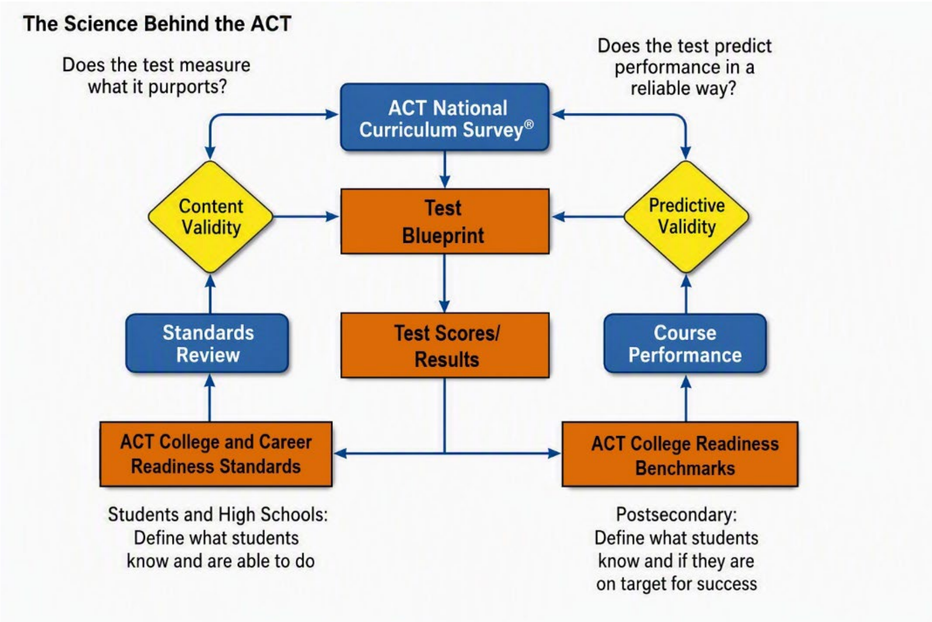
The first step in developing the ACT was to synthesize research on high-value skill targets, those that can be shown to offer the most useful evidence of college and career readiness. This evidence was obtained by organizing the knowledge and skills identified by educators and contained in educational standards into the assessment content framework.

The next step was to use this research to develop content specifications and task models that articulated the evidence needed to monitor student progress. Tasks were then generated from these specifications and assembled into test forms based on test blueprints.

The test blueprints specify constraints on various factors, including, but not limited to, content coverage, item difficulty, cognitive complexity, reading load, and the time required for an item. Test forms are then administered, and student performance data are collected.

Figure 1.1 helps illustrate how a validity argument is composed of multiple sources of research, empirical data, and other forms of evidence. Content validity is shown to be based on research. Predictive validity information flows in primarily from the ACT and, to a lesser extent, the ACT® WorkKeys® assessments. Both channels supply information about which knowledge and skills are needed to perform well on the ACT, thus supporting an iterative model of refinement that serves the common goal of determining whether a student is college and career ready.

**Figure 1.1.** The Full Picture: Evidence and Validity




## 1.5 ACT's Commitment to Fair Testing

Fairness is an essential quality of testing and is related to issues such as testing experience, possible measurement bias, equitable score interpretations, and accessibility (students' ability to accurately demonstrate the extent of their knowledge and skills). Since publication of the original edition in 1988, ACT has endorsed the *Code of Fair Testing Practices in Education* (*Code*; Joint Committee on Testing Practices, 2004), a statement of the obligations to test takers of those who develop, administer, or use educational tests and test data. The development of the *Code* was sponsored by a joint committee including the American Counseling Association, the American Educational Research Association, the American Psychological Association, the American Speech-Language-Hearing Association, the National Association of School Psychologists, the National Association of Test Directors, and the National Council on Measurement in Education, to advance, in the public interest, the quality of testing practices.

The *Code* sets forth fairness criteria in four areas: developing and selecting appropriate tests, administering and scoring tests, reporting and interpreting test results, and communicating with test takers. Separate standards are provided for test developers and test users in each of these four areas. According to the *Code*, for example, test developers should provide "tests that are fair to all test takers regardless of age, gender, disability, race, ethnicity, national origin, religion, sexual orientation, linguistic background, or other personal characteristics" (p. 2). Test developers should "avoid potentially offensive content or language" (p. 4) and "evaluate the evidence to ensure that differences in performance are related to the skills being assessed"

(p. 4). ACT's endorsement of the *Code* represents a commitment to vigorously safeguarding the rights of individuals participating in our testing programs.

Similarly, ACT endorses and is committed to complying with the *Code of Professional Responsibilities in Educational Measurement* (NCME Ad Hoc Committee on the Development of a Code of Ethics, 2023), a statement of professional responsibilities for those who develop assessments; market and sell assessments; select assessments; administer assessments; interpret, use, and communicate assessment results; educate about assessments; evaluate assessment programs; and conduct research on assessments. One of those responsibilities is to "develop assessment products and services that elicit evidence of the measured constructs similarly for all examinees regardless of characteristics such as gender, ethnicity, race, socioeconomic status, disability, religion, age, or national origin" (Section 1.2).

Ensuring fairness in a test is a critically important goal. Lack of fairness must be prevented or detected and eliminated at all stages of test development, test administration, and scoring. The work of ensuring fairness starts with the design of the test and test specifications. It then continues through every stage of the test development process, including item writing and review, item pretesting, item selection and forms construction, and forms review. ACT makes every effort to see that ACT tests are fair to the populations for which the tests are intended and is committed to participating in ongoing dialogues about assessment fairness.

## 1.6 The Population Served by the ACT

Over two million students take the ACT each year. Around 3,000 postsecondary institutions (including scholarship agencies, state educational systems, individual public and private universities, 4-year colleges, junior and community colleges, nursing schools, and technical schools) require or recommend that applicants submit ACT test results.

For the majority of students, postsecondary education begins shortly after high school graduation. Students typically take the ACT during their sophomore, junior, or senior year of high school or shortly before they enroll at a postsecondary institution. Thus, most students who take the ACT are between the ages of 16 and 20.

Most analyses reported in this edition of the *ACT Technical Manual* are based on approximately 300,000 records from eight ACT forms administered from June 2025 to April 2026 to large samples during Saturday ACT National testing events. Self-reported data describing this ACT examinee sample are summarized in Table 1.1. Note that the broader ACT testing population in the United States has somewhat different demographic characteristics because it also includes students who took the ACT during the school day through ACT State and District testing.

ACT advises students to take the ACT after they have completed a substantial portion of the coursework covered by its tests. Given the curriculum of most secondary schools and the course of study followed by the majority of students, this point is usually reached by spring of junior year. However, this varies from student to student and between the four academic areas measured by the ACT, so many students continue to take the ACT into senior year.

**Table 1.1.** Demographic Characteristics of National ACT Testers, June 2025 to April 2026

| Demographic | | Percentage (%) |
|---|---|---|
| **Gender** | Female | 56.7 |
| | Male | 42.3 |
| | Prefer not to respond | 1.0 |
| **Grade level when tested** | Below 9th grade | 0.3 |
| | 9th grade | 0.7 |
| | 10th grade | 4.9 |
| | 11th grade | 44.0 |
| | 12th grade | 47.9 |
| | Other | 2.2 |
| **Racial/ethnic background** | American Indian / Alaska Native | 0.6 |
| | Asian | 8.0 |
| | Black / African American | 11.2 |
| | Hawai’i Native / other Pacific Islander | 0.1 |
| | Hispanic/Latino | 11.6 |
| | White | 58.8 |
| | Two or more races | 4.8 |
| | Prefer not to respond | 4.9 |

*Note*. Due to rounding, some values may not add to exactly 100%.

## 1.7 Test Preparation

Awareness of and exposure to an assessment before taking it is important in order for students to feel comfortable and confident. ACT offers a variety of free and affordable test preparation solutions for students, parents, and educators.

**Free Practice Tests and Study Guides:** ACT’s library of free practice material can be accessed here. This site includes the following:

- **Preparing for the ACT Test.** This resource includes a full-length practice test, test-taking strategies, and information about what to expect on test day. This publication is available in English and Spanish as a free download for teachers, students, parents, and others. It also includes a second free practice test as a companion.

- **ACT Official Online Practice Test.** ACT provides free access to a full-length practice test that simulates an online testing experience. Students may access both timed and untimed practice tests for each test section.

- **Alternate Assessment Format Samples.** Students who will test with alternate formats of the assessment can prepare by practicing with one of our alternate format samples. Braille, large print, audio, and reader’s script formats are available at no cost and contain a full-length practice test.

- **ACT QuizMe:** Free gamified ACT practice in the form of short quizzes for each ACT subject test.

- **ACT Online Prep.** This resource provides students with an interactive test preparation experience that can be accessed anytime online and includes both structured and adaptive paths. It includes personalized learning paths, practice tests with real ACT test questions, and comprehensive content review.

- **ACT Question of the Day.** We post a daily test question to MyACT to provide students with an opportunity for quick daily practice. Students and teachers can opt to receive a weekly email reviewing the questions posted the week before.

- **Powered by Kaplan.** ACT has partnered with Kaplan to publish three official test preparation products:

    - Self-Paced Course: Delivers bite-sized video lessons on demand so students can learn anywhere, anytime.

    - Live Online Classes: Our top-rated teachers show students what to study and how to study during a series of engaging live classes.

    - Tutoring: Led by expert tutors, students learn test content and strategies in these one-on-one online tutoring sessions. Our instructors adapt to the student’s needs and provide each student with personalized attention and recommendations.

# Chapter 2
# The ACT Test Development

## 2.1 Overview

This chapter describes ACT's test development process, including item and form development procedures. The following principles have shaped and will continue to drive ACT's development agenda:

1. Report results in instructionally relevant ways that support clear interpretation within content areas.

2. Maintain reasonable testing times by assessing what research and evidence show to be the most critical factors for success after high school.

3. Leverage technology to enhance student engagement, produce more meaningful results, and share results in a timely fashion.

4. Increase the emphasis on evidence-centered design, implement best practices as they mature, and improve ACT's capabilities to enact the highest-quality design and development processes.

5. Promote STEM (science, technology, engineering, and math) as important to navigating academic and career pathways with a research-backed STEM score based on both the mathematics and science sections.

6. Reflect the research-validated reality that there are multiple dimensions of readiness and success.

As an educational research organization, ACT uses these principles to drive the development and continuous improvement of our education and workplace solutions, as well as the research agenda associated with them, thereby enabling ACT to fulfill its mission of helping all individuals achieve education and workplace success.

This chapter provides brief overviews of the ACT® National Curriculum Survey®, the content and bias review process, and the statistical criteria for selecting operational items and assembling forms. This chapter concludes with a high-level explanation of ACT scoring procedures, including descriptions of additional scores and indicators.

The test development procedures described in this chapter reflect the processes used to build the enhanced ACT forms that have been administered since 2025. Where relevant, this and subsequent chapters include information from legacy ACT forms to illustrate changes introduced as part of the enhancements and to document continuity in test development practices and score interpretation.

## 2.2 Description of the ACT Tests

The ACT® test consists of a standard battery of three subject tests—English, mathematics, and reading—and optional science and writing tests. These tests measure important content, skills, and concepts taught in high school and necessary for success in college and career. The content specifications describing the knowledge and skills to be measured by the ACT were determined through a detailed analysis of relevant information. ACT uses direct feedback from current high school and postsecondary teachers (via the ACT National Curriculum Survey, as well as through external review of test items) and student data from the ACT and from grades earned in postsecondary courses to verify that the ACT measures knowledge and skills empirically linked to postsecondary and career success. The ACT National Curriculum Survey is described in the subsequent section of this chapter. Information about the specific knowledge and skills measured by each test is provided in Chapter 3. Chapter 7 describes sources of validity evidence supporting the interpretation of ACT scores.

## 2.3 The ACT National Curriculum Survey

The ACT National Curriculum Survey is a one-of-a-kind nationwide survey of educational practices and college and career readiness expectations (ACT, 2007, 2009, 2013a, 2016a, 2020a). Conducted by ACT every few years, the National Curriculum Survey embodies ACT's commitment to ensuring not only that the assessments are consistently valid and relevant but also that they provide information enabling students and workers to be fully ready to embark successfully on rewarding college and career journeys. The most recent survey data were collected during the 2025–26 school year, and a new report will be released during the 2026–27 school year.

ACT surveys thousands of K–12 teachers and college instructors in English and writing, mathematics, reading, and science, as well as a national cross section of workforce supervisors and employees, to determine which skills and knowledge in these subjects are currently taught at each grade level and which skills and knowledge are currently considered essential aspects of college and career readiness.

Questions are also included about which skills from the ACT® Holistic Framework®—a research-based framework that integrates behavioral skills, education and career navigation skills, core academic skills, and cross-cutting capabilities (such as teamwork and critical thinking)—are most integral to college and career success.

ACT uses the results of the National Curriculum Survey to guide the development of ACT assessment solutions, including the ACT test, PreACT®, and ACT® WorkKeys®. ACT conducts the survey to ensure that its assessments are measuring the knowledge and skills that instructors of credit-bearing first-year college courses identify as important for success in each content area or that workforce supervisors identify as important for success on the job and readiness for targeted workforce training.

ACT makes the results of each National Curriculum Survey public to help education and workforce stakeholders make more informed decisions about the skills needed to be successful in postsecondary education and the workplace.

### *2.3.1 The Purpose of the ACT National Curriculum Survey*

The National Curriculum Survey is a crucial step in the process of building and regularly updating a suite of ACT assessments that is empirically aligned to college readiness standards. Survey results help address a critical question: Does the test measure knowledge and skills currently relevant to college and career success? Ultimately, the survey data inform the blueprints for the assessments. Subsequently, results from the assessments are used to validate the ACT® College and Career Readiness Standards and the ACT® College Readiness Benchmarks.

Equally important is predictive validity. Using postsecondary course performance data, ACT answers a second critical question: Does the test accurately predict postsecondary performance? Constant monitoring allows ACT to ensure that the answer to both questions is yes.

ACT uses the findings from the National Curriculum Survey to monitor the test blueprints. This process ensures that the assessments measure not only what is being taught in schools around the country but also what demonstrably matters most for college and career readiness. To maintain relevancy and currency, it is important that assessments be built upon up-to-date evidence of what knowledge and skills matter most according to the assessment context and purpose.

The science behind ACT assessments—that is, the evidence base and ongoing research—is critical to answering the key question of what matters most for college and career readiness. The National Curriculum Survey represents ACT's commitment to

- using evidence and research to develop and validate ACT standards, assessments, and benchmarks;
- maintaining a robust research agenda to report on key educational metrics; and
- developing assessments, reports, and interventions that will help individuals navigate their personal path to success along the kindergarten-through-career continuum.

### *2.3.2 Survey Sample and Process*

For the 2020 ACT National Curriculum Survey, ACT recruited participants via various print and electronic methods (e.g., advertisements, email, social media) and invited participation from educators at the early elementary school, late elementary school, middle school, high school, and college levels who teach courses in English and writing, mathematics, reading (including English language arts and social studies), and science (including biology, chemistry, physics, and Earth and space science) in public and private institutions across the United States. ACT also invited participation from supervisors and employees at a large variety of businesses. Table 2.1 gives the number of survey respondents in each area.

**Table 2.1.** ACT National Curriculum Survey 2020 Respondents

| Area | Number of respondents |
|---|---|
| Early elementary school | 1,214 |
| Late elementary school | 1,213 |
| Middle school | 1,623 |
| High school | 1,619 |
| K–12 administrators | 405 |
| College instructors | 2,883 |
| Workforce supervisors | 405 |
| Workforce employees | 406 |
| Total | 9,768 |

Education participants were asked to rate discrete content knowledge and skills with respect to how important each was to student success in a certain content area. Specifically, K–12 teachers were asked to rate the importance of content knowledge and skills in a given class they taught, while college instructors were asked to rate the importance of content knowledge and skills as prerequisites to success in a given class they taught.

ACT also asked the K–12 teachers to indicate whether they taught particular content knowledge and skills and, if so, whether such knowledge and skills were taught as standard parts of their courses or as part of a review of materials that should have been learned earlier. Some education participants were also asked other content-related questions depending on the grade level they taught.

Workforce participants were asked to rate discrete skills with respect to how important each was to success in entry-level positions. ACT also asked workforce participants to indicate how often employees in their workplace used each of these skills on the job.

Finally, ACT asked all participants questions relevant to current education policy issues (e.g., assessments, technology, standards, student characteristics, and obstacles to success). All results are discussed in the report for the ACT National Curriculum Survey 2020 (ACT, 2020a). To ensure that no single content area would have more influence than another on results, the educational-level totals were averaged across English language arts, mathematics, and science.

## 2.4 Test Development Procedures

### *2.4.1 Test Specifications Overview*

As described below, two major types of test specifications are used in developing ACT tests: content specifications and statistical specifications. Several other factors are considered when new test forms are created, such as meeting passage and item word count requirements, avoiding very long strings of the same response option, and preventing extreme imbalance in the distribution of response options.

**Content specifications.** Content specifications for ACT tests were developed through the curricular analysis discussed above. Those specifications define the approximate number of items from each reporting category and cognitive complexity level on a test form. They also set expectations for diverse representation in passages in terms of gender, ethnicity, region, and community type (urban or rural). To support validity and fairness, ACT ensures that the content specifications include only knowledge and skills aligned to the intended purposes of the test. To include anything else would invite construct-irrelevant variance that could unfairly impact students' scores. While care is taken to ensure that the basic structure of each ACT test remains the same from year to year, the specific characteristics of the test items used in each specification category are reviewed regularly. While the general content of the test remains constant, the particular kinds of items in a specification category may change slightly. The basic content structure of each ACT test is provided in Chapter 3.

**Statistical specifications.** Statistical specifications define target levels of difficulty and discrimination at both the item and form levels to ensure that assembled forms are psychometrically comparable and appropriate for the ACT population. Each test section is designed to meet a specified target for average form difficulty while simultaneously satisfying constraints on the distribution of item characteristics.

For all test sections, acceptable item difficulty (proportion correct) ranges from approximately 0.100 to 0.899. In addition to this overall range, the test specifications require a balanced distribution of items across defined difficulty bands (0.100–0.299, 0.300–0.499, and so forth) to ensure that each test form includes an appropriate mix of low-, moderate-, and high-difficulty items. Such a balance supports measurement precision across the full score scale and enhances the reliability of scores for examinees with differing levels of achievement.

The statistical specifications also constrain item discrimination to support reliable differentiation between examinees. Discrimination is evaluated using point-biserial correlations between item performance and total test scores within the relevant content domain. For the enhanced ACT tests, items are expected to exhibit point-biserial correlations of 0.200 or higher, with only a very small proportion of items permitted in the 0.101–0.199 range. These constraints ensure that the majority of items meaningfully contribute to distinguishing between varying levels of student performance, thereby strengthening overall test reliability.

During test assembly, statistical specifications are applied jointly with content specifications. This integrated approach ensures that each test form simultaneously satisfies psychometric targets (e.g., difficulty and discrimination distributions) and content requirements (e.g., blueprint coverage and balance). Together, these constraints support the development of parallel test forms that are comparable in difficulty, precision, and interpretability.

### *2.4.2 Item Writers*

ACT relies primarily on internal content specialists to develop items. Content specialists are subject matter experts, trained in the disciplines for which they write items. Most have experience teaching at various levels, from high school to university, and at a variety of institutions, from small private schools to large public institutions. ACT makes every attempt to include item writers who represent the diversity of the United States population with respect to ethnic background, gender, and geographic location.

Each content specialist is familiar with an item writer's guide that is specific to the specialist's content area. The guides include example items, test specifications, and ACT's requirements for content and style. Also included are specifications for the fair portrayal of all groups, which includes avoidance of subject matter that may be unfamiliar to members of certain groups within society, a balanced representation of race/ethnicity, and gender-neutral language. Item development assignments are balanced among content specialists to ensure a diversity of material.

Depending on development needs, ACT may contract with external item writers or use automated item generation. Externally contracted item writers are also specialists in the content areas measured by the test and typically have teaching experience. Each potential item writer is required to submit a sample set of materials (written using the item writer's guide) for ACT's evaluation. Contracted item writers are held to the same high-quality standards as internal content specialists, and the same attempts to maintain diversity of material and security of the testing program are made. Automated item generation uses models with interchangeable elements based on items that were administered in the past and exhibited desirable statistical properties (e.g., difficulty and discrimination).

### *2.4.3 Item Writing*

Item-writing assignments are driven by the test blueprint and item pool analyses, with the goal of attaining a wide range of high-quality items to elicit evidence of the knowledge, skills, and abilities measured in each test. A typical assignment is tied to an evidence-based item template and focuses on one or more skills that the item needs to assess. Included in each template is a set of statements describing what evidence of students' knowledge and skills should be elicited by the item.

Items are constructed in ACT's item authoring system. This system also contains item metadata, information about the item flow through the stages of development, comments from reviewers, and item quality metrics.

All items must be educationally important and psychometrically sound. Many items must be constructed because, even with good writers, many field-tested items fail to meet ACT's standards.

Each item writer submits a set of items in a given content area. All mathematics items developed recently are discrete (not passage based); some older items belong to a set (i.e., several items are based on the same paragraph or chart). All items on the English and reading tests are related to prose passages. Some reading items may be related to visual or quantitative information, such as graphs and tables, attached to a passage. All items on the science test are related to passages that generally contain data presentations such as graphs and tables.

### *2.4.4 Review of Items*

#### Content Review

After an item (or set of items) is written, it is reviewed by numerous content specialists to verify that it meets all of ACT's standards. It is edited to meet ACT's specifications for content accuracy, word count, item classification, item format, and language. During the review and editing process, all test materials are reviewed for fair portrayal and balanced representation of groups within society and for gender-neutral language.

After internal item reviews are completed, ACT invites external reviewers with knowledge and experience in those content areas, including practicing secondary and postsecondary educators, to refine the items and verify that they should elicit evidence of the intended constructs. During external review, every item is independently reviewed by four to six subject matter experts from across the United States, each of whom has extensive experience with students at or around the grade levels at which the test content is typically taught. These reviewers evaluate the items for content accuracy, item format, and effectiveness of language in terms of leveling, precision, and fairness.

#### Fairness Reviews

Fairness reviews play an essential role in the development of ACT assessments. To help ensure that content is fair, unbiased, and accessible, ACT conducts external fairness reviews for all items before field-testing and for entire test forms before they become operational. In this context, "accessible" means that examinees can access the construct measured by the assessment and accurately demonstrate their construct-relevant knowledge and skills when responding to test items. Avoiding content that is potentially biased is one important aspect of accessibility. Chapter 4 describes ACT's approach to another aspect of accessibility: designing tests and providing testing accommodations for English learners and students with disabilities.

The external fairness review panel consists of experts in diverse areas of education who have experience working with diverse populations. Passages and items are reviewed to help verify that content is not unfair, biased, or insensitive. All comments are reviewed by ACT content specialists, and appropriate changes are made. Because our stakeholders count on national representation to maintain the comparability of test scores, ACT selects reviewers for both content and fairness reviews so that no state is overrepresented.

### *2.4.5 Field-Testing*

ACT field-tests every item before it appears on an operational form to verify that the item functions properly. Items should not be too easy or too difficult, should contribute to precise measurement of the intended construct, and should contain no errors. For the enhanced ACT, each multiple-choice test contains embedded field-test items, which are unscored items that do not contribute to the examinees' scores. These items are administered within the operational section tests under the same standardized testing conditions and time constraints as scored items. The embedded field-test items are indistinguishable from scored items, preserving the natural flow of the assessment and supporting examinee engagement with all items.

After content and fairness reviews, the items and passages that are judged acceptable are assembled into field-test units or item sets appropriate to the structure of each section test. For the English, reading, and science tests, the field-test units include a passage or stimulus with its associated items. For mathematics, field-test items are inserted throughout the test as discrete items. Field-testing for the writing test is conducted as a separate administration, with writing field-test booklets randomly spiraled across the examinee sample.

Embedded field-testing allows ACT to collect field-test data as part of the regular testing experience and allows data from each student to contribute to the item pools for multiple content areas. This approach also aligns the testing experience across school-day, state or district, and national test administrations and creates an inclusive sample for testing new items.

### Analysis of Field-Test Items

After administration, ACT conducts item analyses to examine the psychometric properties of field-test items and flag potentially problematic items. These analyses include item difficulty, item discrimination, distractor analyses, and differential item functioning for gender and ethnicity.

Examinee performance is evaluated in relation to performance on scored operational items in the corresponding ACT content area. For each item, the sample is divided into low-, medium-, and high-performing groups based on examinees' scores on the corresponding ACT section test. The cutoff scores for the three groups are the 27th and 73rd percentiles in the distribution of those scores. These percentiles maximize the critical ratio of the difference between the mean scores of the upper and lower groups, assuming that the standard error of measurement in each group is the same and that the scores for the entire examinee population are normally distributed (Millman & Greene, 1989). Proportions of students in each of the groups correctly answering each field-test item are tabulated, as are the proportions in each group who select each of the incorrect options. The biserial and point-biserial correlation coefficients of each field-test item are also computed.

Items that are either too difficult or too easy are eliminated or revised for future field-testing, as are items that fail to discriminate between students of high and low educational achievement (as measured by their corresponding ACT section test scores). The biserial and point-biserial correlation coefficients, as well as the differences between proportions of students answering the item correctly in each of the three groups, are used as indices of the discriminating power of the field-test items.

Differential item functioning (DIF) analyses are also conducted. DIF is a statistically significant difference between the odds of a certain group (the focal group) answering the item correctly and the odds of a comparison group (the reference group) answering the item correctly when students in the two groups have similar levels of achievement with respect to the content being tested. The procedure currently used for the analysis is the Mantel-Haenszel common odds ratio procedure (MH). Items exhibiting DIF that is large in magnitude and statistically significant are examined by a diverse panel of external fairness reviewers, who evaluate whether there is a content-based explanation for the DIF.

Each item is reviewed following the item analysis. ACT staff members scrutinize items flagged for statistical reasons or DIF to identify possible problems. In some cases, items may be revised and field-tested again before being considered for operational use. The review process also provides feedback that helps ACT improve the quality of future items.

### *2.4.6 Assembly of New Forms*

Items that are judged acceptable in the review process following field-testing are placed in an item pool. Preliminary ACT forms are constructed using items from this pool; items are chosen that match the content and statistical specifications (described in Chapter 3).

Table 2.2 displays the distributions of item difficulty levels for eight forms administered from June 2025 to April 2026. It also reports mean point-biserial correlations and completion rates. Table 2.2 indicates that the ACT forms included a small number of items with *p*-values falling outside the desired range of 0.100–0.899 for English, mathematics, reading, and science. Such items were slightly easier than expected based on data from field-testing.

The completion rate is an indication of whether a test is speeded for a group of students. A test is considered speeded if many students do not have sufficient time to answer the items. The completion rate reported in Table 2.2 for each test is the average completion rate for eight national test forms administered from June 2025 to April 2026. The completion rate for each test is computed as the average percentage of examinees who answered all of the last five items.

**Table 2.2.** Difficulty[a] Distributions and Mean Discrimination[b] Indices for ACT Test Items, June 2025 to April 2026

| Statistic | | English | Mathematics | Reading | Science |
|---|---|---|---|---|---|
| Difficulty range | 0.100–0.299 | 2.19% | 11.28% | 2.78% | 3.68% |
| | 0.300–0.499 | 19.06% | 29.88% | 18.06% | 19.85% |
| | 0.500–0.699 | 44.69% | 35.06% | 37.50% | 41.18% |
| | 0.700–0.899 | 32.50% | 23.48% | 39.81% | 33.82% |
| | 0.900–1.000 | 1.56% | 0.30% | 1.85% | 1.47% |
| No. items[c] | | 320 | 328 | 216 | 272 |
| Mean difficulty | | 0.62 | 0.55 | 0.64 | 0.62 |
| Mean discrimination | | 0.44 | 0.42 | 0.43 | 0.41 |
| Mean completion rate | | 94% | 95% | 97% | 96% |

[a] Item difficulty is the proportion of examinees who correctly answered the item. [b] Item discrimination is the point-biserial correlation coefficient, also known as the item-total

correlation. [c] Each test form consists of 40 scored items for English, 41 for mathematics, 27 for reading, and 34 for science.

### 2.4.7 Content and Fairness Review of Test Forms

The preliminary versions of the test forms are subjected to several reviews to ensure item quality and confirm that the overall test forms meet content and statistical specifications and exemplify best practices supporting fair and accessible testing. ACT staff performs the first review. Items are checked for content accuracy and conformity to ACT style. The items are also examined for clues that could allow test-wise students to answer the items correctly even though they lack the required subject-area knowledge or skills. All ACT test forms also go through an external content review. Each form is reviewed by four to six educators from around the United States, each of whom has extensive experience with students at or around the grade levels at which the test content is typically taught. These reviews follow a process similar to that used during the item development external content review. In addition to focusing on individual items, however, the reviewers also consider the quality of the form as a whole. They judge the form's distributions of content and cognitive complexity to make sure that there is no over- or underrepresentation in any category. Reviewers also look for cluing between items and other issues that could lessen the usefulness of the resulting scores.

Additionally, all newly developed ACT forms must go through external fairness reviews to support fair, equitable, and inclusive assessments that are accessible to all regardless of differences in background or perspective. As with the earlier fairness review, reviewers are experts in diverse areas of education and have experience working with diverse populations. At this stage, reviewers examine individual items and passages, but they also consider the preliminary form as a whole. Forms should be balanced in multicultural and gender representation, and while it is impossible, given the limited amount of material in each test form, to represent every group in every form, a good-faith effort to represent diversity should be discernable.

After the external reviews are complete, ACT summarizes the results. All comments from the consultants are reviewed by ACT content specialists, and appropriate changes are made to the test forms. Whenever significant changes are made, items and/or passages are replaced, and the forms are again reviewed by the appropriate consultants and by ACT staff. If no further changes are needed, the test forms are prepared for publishing.

### 2.4.8 Review Following Operational Administration

After each operational administration, item analysis results are reviewed for any anomalies, such as substantial changes in item difficulty and discrimination indices between the field-test and operational administrations. Only after all anomalies have been thoroughly checked and the final scoring key approved are score reports produced. Examinees may challenge any items they feel are questionable. Once a challenge is raised and reported, the item is reviewed by content specialists in the content area assessed by the item. In the event that a problem is found with an item, actions are taken to eliminate the influence of the problem item as necessary and appropriate. In all cases, each person who challenges an item is sent a response indicating the results of the review.

After each operational administration, differential item functioning (DIF) analysis is also conducted on the test data. As with the field-testing analysis, ACT uses the Mantel-Haenszel common odds ratio procedure (MH) to identify evidence of potential item bias. Items with MH statistics exceeding certain tolerance levels—determined based on established criteria—are flagged. The flagged items are reviewed by content specialists to identify possible content-based explanations for the MH results. If a problem is confirmed, ACT takes appropriate action to eliminate the influence of the problem item.

Table 2.3 lists the percentages of ACT items that exhibited DIF according to the MH procedure for eight forms administered from June 2025 to April 2026. Analyses were conducted to compare item performance for female and male students as well as racial/ethnic groups. Table 2.3 indicates which group was favored by the DIF, meaning the group performed better than expected on the item when we controlled for overall test performance. Note that although DIF is statistical evidence that an item may be biased, approximately 5% of items are expected to be flagged even when there is truly no DIF. In general, DIF flagging rates were near or below the expected 5%. For example, in the gender DIF analysis, the percentages of items flagged ranged from 2.2% in science to 6.7% in mathematics. In the race/ethnicity analysis, the percentage of items flagged for Asian–White DIF sometimes exceeded the expected 5% rate (13.8% in English, 8.8% in mathematics). These values were in the 2%–7% range in the past, so they will be monitored in coming years to determine whether this is an anomaly or a trend.

**Table 2.3.** ACT Test Items Exhibiting Differential Item Functioning, June 2025 to April 2026

| Comparison | Favored group | English | Mathematics | Reading | Science |
|---|---|---|---|---|---|
| Female–Male | Female | 0.6% | — | 3.2% | — |
| | Male | 1.3% | 6.7% | 0.5% | 2.2% |
| American Indian/Alaska Native–White | American Indian/ Alaska Native | — | — | — | — |
| | White | — | — | — | 0.4% |
| Asian–White | Asian | 9.7% | 6.1% | 2.3% | 2.9% |
| | White | 4.1% | 2.7% | — | — |
| Black/African American–White | Black/African American | — | — | — | — |
| | White | 1.9% | 8.2% | 3.7% | 2.2% |
| Hawaii Native/ Other Pacific Islander–White | Hawaii Native/ Other Pacific Islander | — | — | — | — |
| | White | — | — | — | — |
| Hispanic/Latino–White | Hispanic/Latino | — | — | — | — |
| | White | — | 0.3% | — | — |

*Note.* Total item counts were 320 for English, 328 for mathematics, 216 for reading, and 272 for science.

## 2.5 Test Development Procedures for the Writing Test

This section describes the procedures for developing essay prompts for the ACT writing test. These include many of the same steps used to develop the multiple-choice tests.

### *2.5.1 Prompt Writers*

ACT writing prompts are produced by internal content specialists. ACT writing specialists have broad professional experience in secondary and postsecondary classrooms and in the field of writing assessment.

### *2.5.2 Prompt Construction*

Prompts developed for the writing test provide topics with enough complexity and depth that examinees can write thoughtful and engaging essays. Topics are carefully chosen so that they are neither too vast nor too simplistic and do not require specialized prior knowledge. In constructing prompts, ACT writing specialists take into account that a student must be able to respond within the 40-minute time limit.

### *2.5.3 Content and Fairness Review of Prompts*

After writing test prompts are developed and refined by ACT writing specialists, the prompts go through a rigorous review process with external experts. These fairness and bias experts carefully review each prompt to ensure that neither the language nor the content of a prompt will be offensive to a test taker and that no prompt will disadvantage any student from any geographic, socioeconomic, or cultural background. Reviewers also help ensure that prompts are accessible and engaging to students by evaluating prompt content in relation to student knowledge, experience, and interests.

### *2.5.4 Field-Testing of Prompts*

ACT periodically conducts a special field-test study to evaluate new ACT writing prompts and select those suitable for operational use. Students from across the United States—from rural and urban settings, small and large schools, and public and private schools—write responses to the new prompts, which are then read and scored by ACT-trained readers.

Prompts are evaluated from both content and statistical perspectives to ensure that scores (reported on a scale of 2 to 12) are comparable across different test forms and different administrations. In each field-test study, anchor prompts and new prompts are administered to randomly equivalent groups of approximately 1,000 students per prompt.

Each student takes two prompts, and the order in which the prompts are taken is counterbalanced. Prompts are spiraled within classrooms so that, across all participating students, randomly equivalent groups of students take each prompt, with about half of the students taking a prompt first and the rest taking it second.

### 2.5.5 Review of Field-Tests and Operational Administration

Once scoring of the new writing test prompts has been completed, the prompts are statistically analyzed to judge their acceptability. ACT applies the acceptability criteria after examining the relationships among scores on newly field-tested prompts and older (anchor) prompts. Specifically, the 2-to-12 score distributions should align, and there should be students scoring at the top of the score scale. Also, equating results should show that equating errors are within expected ranges at all score points, and the raw-to-scale score conversion tables, which are used to generate scores (from 1 to 36) that contribute to the ACT ELA (English language arts) score, should exhibit desirable properties (see Chapter 6.2 for more information about writing equating).

## 2.6 ACT Scores

This section briefly introduces the scores generated from student responses to the ACT test. Chapter 5 provides additional information about these scores and ACT score reports. This section concludes with a summary of ACT policies concerning scoring appeals and inquiries.

### 2.6.1 ACT Scale Scores

For each multiple-choice test section on the ACT (English, mathematics, reading, and science), the raw scores (number of correct multiple-choice responses) are converted to scale scores ranging from 1 to 36. The Composite score is the average of the English, mathematics, and reading scale scores rounded to the nearest whole number (fractions of 0.5 or greater round up). The minimum Composite score is 1; the maximum is 36. See Chapter 6 for more details about the creation and maintenance of the 1-to-36 ACT scales.

For the writing test, scores of 2–12 are reported for each of the four domains of writing: Ideas and Analysis, Development and Support, Organization, and Language Use and Conventions (these are described in detail in Chapter 3). The subject-level writing test score, also 2 to 12, is the rounded average of the four domain scores. Writing scores are converted to a 1-to-36 scale only for the purpose of calculating the ELA score; the 1-to-36 writing scores are not reported. More detailed information about how the writing test is scored can be found in Chapter 5.

### 2.6.2 STEM and ELA Scores

Since fall 2015, ACT has reported a STEM score, which is the average of the 1-to-36 mathematics and science scale scores rounded to the nearest integer (fractions of 0.5 or greater round up). Only students who receive scores on the mathematics and science tests receive an ACT STEM score.

In fall 2015, ACT also began reporting a combined ELA score. The ACT ELA score is the rounded average of the English, reading, and writing scores. Only students who take all three of these tests can receive an ELA score. For the calculation of ELA scores, the sum of the writing domain scores is converted to a scale of 1 to 36. However, this 1-to-36 writing scale score is not reported independently. Procedures for obtaining the 1-to-36 writing scale scores are described in Chapter 6.

### 2.6.3 Reporting Category Scores and Readiness Ranges

English, mathematics, reading, and science items align with reporting categories linked to the ACT College and Career Readiness Standards and other standards that target college and career readiness. There are three reporting categories each for English, reading, and science and eight for mathematics. Students receive a score in each reporting category, and score reports show corresponding Readiness Ranges, which indicate the range of scores expected of students who met or exceeded the ACT College Readiness Benchmark in that content area. The ACT Readiness Ranges appear on the Student Score Report and the High School Score Report. The combination of reporting category scores and the ACT Readiness Ranges provides educators and students with information that more clearly shows where students require the most assistance. Descriptions of the reporting categories are provided in Chapter 3.

### 2.6.4 Progress Toward the ACT WorkKeys National Career Readiness Certificate (NCRC) Indicator

The Progress Toward the ACT® WorkKeys® National Career Readiness Certificate™ (NCRC®) Indicator is based on students' ACT Composite scores. It provides an estimate of students' most likely performance on the three assessments that lead to the WorkKeys NCRC, an assessment-based credential that certifies foundational work skills important for job success across industries and occupations. The NCRC is based on the results of three WorkKeys assessments: ACT® WorkKeys® Applied Math, ACT® WorkKeys® Graphic Literacy, and ACT® WorkKeys® Workplace Documents. Scores on these assessments determine an individual's certificate level—no certificate, Bronze, Silver, Gold, or Platinum. The NCRC gives individuals evidence that they possess the skills employers deem essential to workplace success. More information about the NCRC can be found here.

### 2.6.5 Scoring Verifications and Inquiries

Electronic scanning devices are used to score the four multiple-choice tests on the ACT, thus minimizing the potential for scoring errors. If a student believes that a scoring error has been made, ACT hand-scores the answer document (for a fee) upon receipt of a written request from the student. Strict confidentiality of each student's record is maintained.

If a student believes that a writing test essay has been incorrectly scored, that score may be appealed. ACT will verify (for a fee) that the essay was scored by at least two independent, qualified readers—one of which may have been ACT's automated scoring engine, CRASE®—and by a third reader in the event that the two scores differed by more than 1 point in any domain. ACT will also verify that the essay was properly captured and displayed to readers. In the case of an online test that was scored only by the automated scoring engine, ACT will have an independent human rater validate the score. If errors are discovered during score verification, ACT will rescore the essay and refund the score verification fee.

For certain test dates (found on ACT's website), examinees may obtain (for a fee) a copy of the test items used in determining their scores, a list of the correct answers, a list of their answers, and a table to convert raw scores to the reported scale scores. (For an additional fee, a student may also obtain a copy of his or her answer document.) These materials are available only to students who test during regular administrations of the ACT on specified national test dates. If for any reason ACT must replace the test form scheduled for use at a test center, this offer is withdrawn, and the student's fee for this optional service is refunded.

ACT reserves the right to cancel test scores when there is reason to believe the scores are invalid. Cases of irregularities in the test administration process—falsifying one's identity, impersonating another examinee, unusual similarities in the answers of examinees at the same test center, examinee misconduct, or other indicators that the test scores may not accurately reflect the examinee's level of educational achievement—may result in ACT's canceling the test scores. For a detailed description of how ACT handles score cancelations, refer to ACT's Terms and Conditions of Registration.

# Chapter 3
# Content Specifications

## 3.1 Overview

The ACT® test is constructed to meet specifications for content balance within the assessment domains. The content specifications define ranges for the number of items in each content category and at each level of cognitive complexity. The content specifications may also set test-specific requirements for the number of passages, distribution of passage genres, passage and item word counts, and diverse representation in passages in terms of gender, ethnicity, region, and community type (urban or rural). These content blueprints ensure that the knowledge and skills in the content domains are sampled consistently across test forms. The following chapter describes the assessment domain and content blueprint for each of the four multiple-choice ACT tests and the writing test.

## 3.2 English Test

### *3.2.1 Description of the English Test*

The ACT English test has 50 items (40 scored items and 10 field-test items) and a 35-minute time limit. The test puts the student in the position of a writer who is revising and editing a text and measures a student's understanding of the conventions of standard written English (grammar, usage, and mechanics), production of writing (topic development, organization, unity, and cohesion), and knowledge of language (word choice, style, and tone). The test consists of multiple passages, each accompanied by a sequence of multiple-choice test items. Different passage types—informational, argumentative, and narrative—are employed to provide a variety of rhetorical situations. Students must use the rich context of the passages to make editorial choices, demonstrating their understanding of writing strategies and conventions. Passage length also varies, with shorter passages featuring 5 items and longer passages featuring 10 items. Passages are written not only to appropriately assess writing and language skills but also to reflect students' interests and experiences. Spelling and the rote recall of grammar rules are not tested.

#### Cognitive Complexity and Depth of Knowledge (DOK)

DOK (Webb, 2002) is a rough-grained, judgment-based measure of a test item's cognitive complexity that is used in many educational contexts. The ACT English test assesses skills that vary in cognitive complexity using items at DOK Levels 1, 2, and 3. All English items are classified by ACT content experts according to the level descriptions in Table 3.1.

**Table 3.1.** DOK Level Descriptions for English

| Depth of Knowledge level | Description |
|---|---|
| DOK1 | Requires the recall of information, such as a fact, term, definition, or simple procedure. |
| DOK2 | Requires mental processing that goes beyond recalling or reproducing an answer. Students must make some decisions about how to approach a problem. |
| DOK3 | Requires planning, thinking, explaining, justifying, using evidence, conjecturing, and postulating. |

### *3.2.2 English Scores and Reporting Categories*

Four scores are reported for the ACT English test: a total test score based on all 40 scored items, plus three reporting category scores. The total test score is reported on the ACT English scale, which ranges from 1 to 36. That score is averaged with the reading and writing test scores to determine the ELA score (see Chapter 5 for more information about the derivation of the ELA score). The three reporting categories associated with the English test are Production of Writing, Knowledge of Language, and Conventions of Standard English. A brief description of the reporting categories is given below. ACT score reports provide the percentage of correctly answered items in each reporting category and a Readiness Range indicating the range of scores expected of students who meet the ACT® College Readiness Benchmark for English (18).

#### Production of Writing

Students apply their understanding of the rhetorical purpose and focus of a piece of writing to develop a topic effectively. They use various strategies to achieve logical organization, topical unity, and cohesion.

#### Knowledge of Language

Students demonstrate effective language use by ensuring precision and concision in word choice and maintaining consistency in style and tone.

#### Conventions of Standard English

Students apply their understanding of the conventions of Standard English grammar, usage, and mechanics to revise and edit text.

### *3.2.3 English Test Blueprints*

Table 3.2 shows the current target distribution of test items across reporting categories on each ACT English test form.

**Table 3.2.** Specification Ranges by Reporting Category for English (Scored Items Only)

| Reporting category | Number of items | Percentage of test |
|---|---|---|
| Production of Writing | 15–17 | 38%–43% |
| Knowledge of Language | 7–9 | 18%–23% |
| Conventions of Standard English | 15–17 | 38%–43% |
| Total | 40 | 100% |

## 3.3 Mathematics Test

### *3.3.1 Description of the Mathematics Test*

The ACT mathematics test has 45 items (41 scored items and 4 field-test items) and a 50-minute time limit. The test considers the whole of a student's mathematical development up through topics typically taught at the beginning of Grade 12 in U.S. schools, focusing on prerequisite knowledge and skills important for success in college mathematics courses and career training programs. The domain is divided into Preparing for Higher Mathematics (PHM) and Integrating Essential Skills (IES).

The mathematics construct requires making sense of problems and context, representing relationships mathematically, accessing appropriate mathematical knowledge from memory, incorporating given information, modeling, performing mathematical computations and manipulations, interpreting, applying reasoning skills, justifying, making decisions based on mathematics, and appropriately managing the solution process. The test emphasizes quantitative reasoning and application over extensive computation or the memorization of complex formulas. Items focus on what students can do with the mathematics they have learned, which encompasses not only mathematical content but also mathematical practices.

Some degree of computational fluency is required. A calculator is encouraged but not required. Items are designed so that a sophisticated calculator does not provide a significant advantage over a four-function calculator. Items are also designed so that all problems can be done without a calculator in a reasonable amount of time.

Each item has four response options, and students are instructed to choose the correct option. The test contains problems ranging from easy to very challenging so that readiness levels can be reliably reported for students with different degrees of preparedness. Extended accessibility supports provide for fair and comparable mathematics scores across a range of circumstances. More information on accessibility can be found in Chapter 4.

#### Cognitive Complexity and Depth of Knowledge (DOK)

The ACT mathematics test assesses skills that vary in cognitive complexity using items at DOK Levels 1, 2, and 3. DOK is judged relative to well-prepared students in Grades 11–12. All mathematics items are classified by ACT content experts according to the level descriptions in Table 3.3.

**Table 3.3.** DOK Level Descriptions for Mathematics

| Depth of Knowledge level | Description |
|---|---|
| **DOK1** | Requires the recall of information, such as a fact, term, definition, or simple procedure. Requires students to demonstrate a rote response or perform a simple procedure. |
| **DOK2** | Requires mental processing that goes beyond recalling or reproducing an answer. Students must make some decisions about how to approach a problem. |
| **DOK3** | Requires planning, thinking, explaining, justifying, using evidence, conjecturing, and postulating. The cognitive demands are complex and abstract. |

### *3.3.2 Mathematics Scores and Reporting Categories*

Nine scores are reported for the ACT mathematics test: a total test score based on all 41 scored items, plus eight reporting category scores based on different domains of mathematical knowledge, skills, and practices. The total test score is reported on the ACT mathematics scale, which ranges from 1 to 36. That score is averaged with the science score to determine the STEM score, which is related to success in postsecondary science, technology, engineering, and mathematics courses (see Chapter 5 for more information about the derivation of the STEM score).

There are eight mathematics reporting categories designed to give more detail about a student's mathematical achievement. The additional reporting category scores show a pattern of strengths and weaknesses that can differ among students with the same mathematics score. The test is first divided into the Preparing for Higher Mathematics (PHM) and Integrating Essential Skills (IES) reporting categories. The PHM score is then divided into separate scores for Number & Quantity, Algebra, Functions, Geometry, and Statistics & Probability. A crosscutting reporting category, Modeling, draws upon items from all the other categories to give a measure of producing, interpreting, understanding, evaluating, and improving models. Descriptions of each reporting category follow, and Table 3.4 shows the number of items that contribute to each reporting category score. ACT score reports provide the percentage of items in each reporting category answered correctly and a Readiness Range indicating the range of scores expected of students who meet the ACT College Readiness Benchmark for mathematics (22).

#### Preparing for Higher Mathematics

This reporting category captures the mathematics that students have learned more recently. This category is divided into the following five subcategories.

##### *Number & Quantity*

Students demonstrate an understanding of and fluency with rational numbers and the four basic operations, and they work with irrational numbers, including approximating irrational numbers with rational numbers. They use properties of the real number system. They show their knowledge of complex numbers, compute in this system, and work with the properties of complex numbers. They use vectors and matrices and view them as number systems with properties, operations, and applications.

### *Algebra*

Students use expressions to solve problems, and they show an understanding of solving equations. They manipulate and evaluate expressions involving integer exponents, square and cube roots, and scientific notation. They solve linear equations and make connections between their graphs and proportional relationships. They use their understanding of linear equations to make sense of other kinds of equations and inequalities: what their graphs look like, how to solve them, and what kinds of applications they have for modeling. They demonstrate extended proficiency with equations by using quadratic, polynomial, exponential, rational, and radical equations as well as systems of equations. They create expressions, equations, and inequalities to represent problems and constraints. They see rational expressions as systems analogous to rational numbers, apply the binomial theorem, and solve simple matrix equations that represent systems of linear equations.

### *Functions*

Students demonstrate an understanding of what a function is and what its characteristics are. They can define functions using two variables or function notation. They use their understanding of the general properties of functions to solve problems using new functions they create. They determine what the characteristics of functions mean in the context of a problem. They work with functions that have no equation and functions that follow the pattern of an equation. They work with particular families of a variety of functions, such as linear, quadratic, polynomial, radical, exponential, logarithmic, and piecewise, and apply them in modeling situations. They associate changes in the parameters of a function that has a form such as $f(x) + c$ or $f(-ax)$ with transformations of the graph of the function. They compose functions and use inverse functions to solve equations with more than one solution. They connect the trigonometry of right triangles to the unit circle to make trigonometric functions and use them to model periodic behavior. They apply the algebraic properties of trigonometric functions, such as angle addition properties.

### *Geometry*

Students show an understanding of congruence and similarity using transformations like translations and dilations. They investigate the relationships between the angles formed by parallel lines and a transversal. They model and solve problems with geometric objects. Students find values such as the area of a circle and the volume of cylinders, pyramids, and cones. They demonstrate an understanding of trigonometric ratios as functions of angles, and they solve right-triangle problems using these ratios and the Pythagorean theorem. In the coordinate plane, they recognize conditions for parallel and perpendicular lines, split a line segment into pieces with a given ratio of lengths, find areas of plane figures, and develop equations for circles and parabolas. They use trigonometry to compute the area of a triangle, and they apply the law of sines and the law of cosines to answer questions about nonright triangles. They derive equations for ellipses and hyperbolas.

#### *Statistics & Probability*

Students identify patterns of association between two quantities by analyzing scatterplots and two-way tables. They fit linear models to datasets and use the models to solve problems. They demonstrate their knowledge of the role of randomness in sample surveys, experiments, and observational studies. They analyze data to make inferences and evaluate claims. They demonstrate an understanding of statistical independence. They relate the sample space to events defined in terms of "and," "or," and "not," and they calculate probabilities using empirical results, independence assumptions, and the ideas of conditional probability. They understand the multiplicative rule for conditional probability and apply permutations and combinations as tools for counting. They model a sample space with a random variable by giving a numerical value to each event. They apply expected value and probability to help inform their decisions.

### Integrating Essential Skills

This reporting category focuses on whether students can apply their knowledge and skills to solve problems of moderate to high complexity. Topics include rate and percentage; proportional reasoning; units of measure; solving problems with rational numbers; constructing and solving simple equations; scale drawings; cross sections; area, perimeter, circumference, surface area, and volume; simple measures of center and spread; probability; and random sampling.

In addition to learning more content as they progress through their studies, students should also grow in sophistication, accumulating and applying skills in higher-order contexts. Therefore, students should be able to solve problems of increasing complexity, combine skills in longer chains of steps, apply skills in more varied contexts, understand more connections, and increase fluency. In order to assess whether students have had appropriate growth, the items in this reporting category are at DOK Levels 2 and 3.

### Modeling

Modeling uses mathematics to represent, through a model, an analysis of an empirical situation. Models often help us predict or understand the actual. However, sometimes knowledge of the actual helps us understand the model, such as when addition is introduced to students as a model of combining two groups. The Modeling reporting category represents all items that involve producing, interpreting, understanding, evaluating, and improving models. Each modeling item is also counted in the other appropriate reporting categories above. Thus, the Modeling reporting category is an overall measure of how well a student uses modeling skills across mathematical topics.

## *3.3.3 Calculator Policy*

Students are encouraged to bring a calculator they are familiar with and can use fluently. Most four-function, scientific, or graphing calculators are permitted. Built-in computer algebra systems are not allowed because they could interfere with the construct, specifically understanding and implementing solutions to various types of equations and inequalities. Students must remove certain kinds of programs from their calculators. Some calculator features are not allowed or must be turned off for security reasons or to prevent disruptions during testing. Current details are available on the ACT website.

### *3.3.4 Mathematics Test Blueprints*

Table 3.4 shows the current target distribution of scored items across reporting categories on each ACT mathematics test form. Test construction also takes into account coverage and variety within each of the categories. The test includes PHM items at DOK Levels 1, 2, and 3. By contrast, it includes IES items only at DOK Levels 2 and 3. This is because IES topics are more practiced and familiar and because putting these familiar skills to work in higher-complexity tasks is important for college readiness.

**Table 3.4.** Specification Ranges by Reporting Category for Mathematics (Scored Items Only)

| Reporting category | Number of items | Percentage of test |
|---|---|---|
| Preparing for Higher Mathematics | 33 | 80% |
| Number & Quantity | 4–5 | 10%–12% |
| Algebra | 7–8 | 17%–20% |
| Functions | 7–8 | 17%–20% |
| Geometry | 7–8 | 17%–20% |
| Statistics & Probability | 5–6 | 12%–15% |
| Integrating Essential Skills | 8 | 20% |
| Modeling | ≥8 | ≥20% |
| Total | 41 | 100% |

*Note*. Each item reported in Modeling is also reported in either Preparing for Higher Mathematics (and the appropriate subcategory) or in Integrating Essential Skills.

## 3.4 Reading Test

### *3.4.1 Description of the Reading Test*

The ACT reading test has 36 items (27 scored items and 9 field-test items) and a 40-minute time limit. The test measures a student's ability to read closely, reason about texts using evidence, and integrate information from multiple sources. The test comprises four units, each of which includes a passage or a pair of passages accompanied by 9 multiple-choice items. Passages in the reading test include both literary narratives and informational texts from the humanities, natural sciences, and social sciences. Visual and quantitative elements may accompany informational passages and contain additional information related to the passage topic. Passages are representative of the kinds of texts commonly encountered in high school and first-year college courses. Each passage is preceded by a heading that identifies the passage type (Literary Narrative or Informational) and names the author; it may also include important background information that helps in understanding the passage.

Reading items focus on the mutually supportive skills that readers apply when studying written materials across a range of subject areas. Specifically, items ask students to determine main ideas; locate and interpret significant details; understand sequences of events; make comparisons; comprehend cause-effect relationships; determine the meaning of context-dependent words, phrases, and statements; draw generalizations; analyze the author's or narrator's voice or method; analyze claims and evidence in arguments; and integrate information from multiple related texts and from different formats (e.g., graphs, diagrams, tables). Items do not test the rote recall of facts from outside the passage or rules of formal logic, nor do they contain questions about vocabulary that can be answered without referring to the passage context.

### Cognitive Complexity and Depth of Knowledge (DOK)

The ACT reading test assesses skills that vary in cognitive complexity using items at DOK Levels 1, 2, and 3. All reading items are classified by ACT content experts according to the level descriptions in Table 3.5.

**Table 3.5.** DOK Level Descriptions for Reading

| Depth of Knowledge level | Description |
|---|---|
| DOK1 | Requires the recall of information, such as a fact, term, definition, or simple procedure. Requires students to demonstrate a rote response or perform a simple procedure. |
| DOK2 | Requires mental processing that goes beyond recalling or reproducing an answer. Students must make some decisions about how to approach a problem. |
| DOK3 | Requires planning, thinking, explaining, justifying, using evidence, conjecturing, and postulating. The cognitive demands are complex and abstract. |

## *3.4.2 Reading Scores and Reporting Categories*

Four scores are reported for the ACT reading test: a total test score based on all 27 scored items, plus three reporting category scores based on specific knowledge and skills. The total test score is reported on the ACT reading scale, which ranges from 1 to 36. That score is averaged with the English and writing scores to determine the ELA score (see Chapter 5 for more information about the derivation of the ELA score). The three reporting categories addressed in the reading test are Key Ideas & Details, Craft & Structure, and Integration of Knowledge & Ideas. ACT score reports provide the percentage of items in each reporting category answered correctly and a Readiness Range indicating the range of scores expected of students who meet the ACT College Readiness Benchmark for reading (22).

### Key Ideas & Details

Students read texts closely to analyze central ideas and themes; summarize information and ideas accurately; and understand sequential, comparative, and cause-effect relationships. They use evidence from the text to make inferences and draw conclusions.

### Craft & Structure

Students determine word and phrase meanings, analyze word choice, analyze text structure, determine authorial purpose and perspective, and analyze points of view. They interpret the rhetorical effects of authorial decisions and differentiate between various perspectives and sources of information.

### Integration of Knowledge & Ideas

Students analyze authors' claims and how arguments are constructed, evaluate reasoning and evidence, and differentiate between facts and opinions. They evaluate multiple sources of information and use evidence to make connections between related texts or between a text and a related graphic.

### *3.4.3 Reading Test Blueprints*

Table 3.6 shows the current target distribution of test items across reporting categories on each ACT reading test form.

**Table 3.6.** Specification Ranges by Reporting Category for Reading (Scored Items Only)

| Reporting category | Number of items | Percentage of test |
|---|---|---|
| Key Ideas & Details | 12–14 | 44%–52% |
| Craft & Structure | 7–9 | 26%–33% |
| Integration of Knowledge & Ideas | 5–7 | 19%–26% |
| Total | 27 | 100% |

## 3.5 Science Test

### *3.5.1 Description of the Science Test*

The ACT science test has 40 items (34 scored items and 6 field-test items) and a 40-minute time limit. This optional test measures the interpretation, analysis, evaluation, reasoning, and problem-solving skills required in the natural sciences. The content of the science test is drawn from biology, chemistry, physics, and Earth and space science.

Students are assumed to have a minimum of 2 years of high school introductory science, which ACT's National Curriculum Survey has identified as typically 1 year of biology and 1 year of physical science or Earth science. It is thus expected that students have acquired the introductory content of biology, physical science, and Earth science, are familiar with the nature of scientific inquiry, and have been exposed to laboratory investigation.

The test presents several sets of scientific information, each followed by a number of multiple-choice test items. The scientific information is conveyed in one of three formats: data representation (scientific graphs, tables, and diagrams), research summaries (descriptions of one or more related experiments), or conflicting viewpoints (two or more brief theoretical models that address the same scientific phenomenon but conflict with one another).

The test assesses and reports on science knowledge, skills, and practices across three domains: Interpretation of Data, Scientific Investigation, and Evaluating Models and Scientific Arguments with Evidence. The knowledge and skills encompassed in each domain were derived from decades of ACT's empirical data and research on college and career readiness in science. The domains and their skills link with quantitatively determined score ranges for the ACT science test and the ACT College Readiness Benchmark in science, which predicts success in science courses at the postsecondary level.

In addition, some of the ACT science items require students to integrate discipline-specific content knowledge (e.g., knowledge specific to physical science or biology) with science and engineering practices and crosscutting concepts in science to assess three-dimensional scientific sense-making present in high school science standards and necessary for success in first-year college courses.

#### Cognitive Complexity and Depth of Knowledge

The ACT science test assesses skills and practices that vary in cognitive complexity using items at DOK Levels 1, 2, and 3, with almost all items at DOK Levels 2 and 3. ACT science experts have worked with several Webb-based systems adapted for science, but none of those systems quite capture the different dimensions associated with items focused on science skills and practices. Even so, all science items are classified by ACT content experts according to the level descriptions in Table 3.7.

**Table 3.7.** DOK Level Descriptions for Science

| Depth of Knowledge level | Description |
|---|---|
| DOK1 | Requires locating, recalling, and/or reproducing information. |
| DOK2 | Requires processing presented information and applying skills and concepts. Students typically must process one or two cognitive steps. |
| DOK3 | Requires use of higher-order thinking, such as analysis and evaluation, and often requires using evidence to justify reasoning. Students must typically process multiple cognitive steps, and the overall tasks tend to be complex and abstract. |

### *3.5.2 Science Scores and Reporting Categories*

Four scores are reported for the ACT science test: a total test score based on all 34 scored items, plus three reporting category scores based on different domains of scientific knowledge, skills, and practices. The total test score is reported on the ACT science scale, which ranges from 1 to 36. That score is averaged with the mathematics score to determine the STEM score, which is related to success in postsecondary science, technology, engineering, and mathematics courses (see Chapter 5 for more information about the derivation of the STEM score). The three reporting categories addressed in the science test are Interpretation of Data, Scientific Investigation, and Evaluating Models and Scientific Arguments with Evidence. A description of each reporting category is provided below. ACT score reports provide the percentage of items in each reporting category answered correctly and a Readiness Range indicating the range of scores expected of students who meet the ACT College Readiness Benchmark for science (23).

#### Interpretation of Data

Students manipulate and analyze scientific data presented in tables, graphs, and diagrams (e.g., by recognizing trends in data, translating tabular data into graphs, interpolating and extrapolating, and reasoning mathematically).

#### Scientific Investigation

Students understand experimental tools, procedures, and design (e.g., by identifying variables and controls) and compare, extend, and modify experiments (e.g., by predicting the results of additional trials). Students also identify goals, constraints, and potential tradeoffs in engineering and design thinking contexts.

#### Evaluating Models and Scientific Arguments with Evidence

Students judge the validity of scientific information and formulate conclusions and predictions based on that information (e.g., by determining which explanation for a scientific phenomenon is supported by new findings). Students evaluate tradeoffs, success criteria, and constraints in engineering and design thinking contexts.

### *3.5.3 Science Test Blueprints*

Table 3.8 shows the current target distribution of test items across reporting categories on each ACT science test form. Table 3.9 shows the current target distribution of test items across science content areas on each ACT science test form.

**Table 3.8.** Specification Ranges by Reporting Category for Science (Scored Items Only)

| Reporting category | Number of items | Percentage of test |
|---|---|---|
| Interpretation of Data | 13–17 | 38%–50% |
| Scientific Investigation | 6–11 | 18%–32% |
| Evaluating Models and Scientific Arguments with Evidence | 8–13 | 24%–38% |
| Total | 34 | 100% |

**Table 3.9.** Specification Ranges by Science Content Area (Scored Items Only)

| Science content area | Number of passages | Number of items | Percentage of test |
|---|---|---|---|
| Biology | 2 | 10–13 | 29%–38% |
| Chemistry | 1–2 | 5–13 | 15%–38% |
| Physics | 1–2 | 5–13 | 15%–38% |
| Earth & space science | 1–2 | 5–13 | 15%–38% |
| Total | 6 | 34 | 100% |

## 3.6 Writing Test

### *3.6.1 Description of the Writing Test*

The ACT writing test is an optional 40-minute essay test that measures students' writing skills—specifically those skills emphasized in high school English classes and entry-level college composition courses. Scores from the writing test indicate students' ability to think critically about an issue, consider different perspectives on it, and compose an effective argumentative essay.

The test consists of one writing prompt that describes a complex issue and provides three different perspectives on the issue. Students are asked to read the prompt and write an essay in which they develop their own perspective on the issue. The essay must analyze the relationship between their own perspective and one or more other perspectives. Students may adopt one of the perspectives given in the prompt as their own, or they may introduce one that is completely different from those given. Their score will not be affected by the point of view they take on the issue.

#### Cognitive Complexity and Depth of Knowledge (DOK)

The cognitive complexity of the writing test essay task is classified as DOK 3 (Table 3.10).

**Table 3.10.** DOK Level Description for Writing

| Depth of Knowledge level | Description |
|---|---|
| DOK3 | Requires planning, thinking, explaining, justifying, using evidence, conjecturing, and postulating |

### *3.6.2 Writing Scores and Domains*

Students who take the optional writing test receive five scores: a single subject-level writing score and four domain scores. The overall writing score is reported on the ACT writing scale, which ranges from 2 to 12.[1] Taking the writing test does not affect the student's section test scores or Composite score. However, a writing test score, along with the overall English and reading test scores, is needed to produce the ELA score. The overall writing score (after it has been converted to a 1–36 scale) is averaged with the English and reading test scores to determine the ELA score (see Chapter 5 for more information about the derivation of the ELA score).

The four writing domains are Ideas & Analysis, Development & Support, Organization, and Language Use & Conventions. A brief description of the writing domains is given below. Scores on the four domains are each reported on a 2–12 scale, and the overall writing score is the rounded average of the four domain scores. The domain scores are based on an analytic scoring rubric; essays are scored on a scale of 1 to 6 in each of the four domains. Paper test responses are independently scored by two trained human raters. Online test responses are scored either by one trained human rater and CRASE® (ACT's automated scoring engine) or by CRASE only (where eligible and when the assigned score meets an established confidence threshold). If any essays receive two ratings that disagree by more than 1 point, a third rater evaluates the essay and resolves the discrepancy. See Chapter 5 for more information about writing performance scoring and the analytic scoring rubric.

### Ideas & Analysis

Scores in this domain reflect the ability to generate productive ideas and engage critically with multiple perspectives on the given issue. Proficient writers understand the issue they are invited to address, the purpose for writing, and the audience. They generate ideas that respond to the situation.

### Development & Support

Scores in this domain reflect the ability to discuss ideas, offer rationale, and strengthen an argument. Proficient writers explain and explore their ideas, discuss implications, and illustrate ideas through examples. They help the rater understand their thinking about the issue.

### Organization

Scores in this domain reflect the ability to organize ideas with clarity and purpose. Organizational choices are integral to effective writing. Proficient writers arrange their essay in a way that clearly shows the relationships among ideas, and they guide the reader through their discussion.

[1] Students who took the writing test between September 2015 and June 2016 received subject-level writing scores reported on a 1–36 scale rather than subject-level scores reported on the current 2–12 scale. It should also be noted that the current 2–12 subject-level writing scores are not comparable to the 2–12 scores from the former writing test (June 2015 and before). Although both tests measure a student's ability to write an effective argumentative essay, the current test has a new design. Moreover, the current test is scored with an analytic rubric, whereas the former test was scored with a holistic 6-point rubric. The score on the former test was the sum of the two raters' 1–6 scores rather than the rounded average of four 2–12 domain scores.

### Language Use & Conventions

Scores in this domain reflect the ability to use written language to clearly convey ideas. Proficient writers adhere to the conventions of grammar, syntax, word usage, and mechanics. They are also aware of their audience and adjust the style and tone of their writing to communicate effectively.

# Chapter 4
# Test Administration, Test Security, and Accessibility and Accommodations

## 4.1 Test Administration Overview

The ACT® test must be administered in a standardized manner to ensure a fair and equitable testing environment for all examinees. Testing staff must strictly adhere to ACT policies and procedures during test administrations. This chapter provides a brief description of the processes used in the administration of the ACT in both paper and online formats.

### *4.1.1 Administration Windows*

The ACT is administered on predetermined test dates and during established test windows. For the ACT National and International tests, these dates and their registration deadlines are available at www.act.org; the ACT can be administered only on these dates and at these times at a given test center. For ACT State and District testing, stakeholders can choose from predetermined dates and windows during the fall and spring.

### *4.1.2 Testing Modes*

The ACT is administered on paper and online for National testing and online for International testing, though there are exceptions for special testing and certain accommodations. State and District testing sites have the option of administering the test on paper or online. Information about the comparability between the paper and online modes may be found in Chapter 6. In addition to offering standard formats, ACT offers accommodations and English learner (EL) supports for examinees approved for these accessibility supports.

### *4.1.3 Testing Locations*

The ACT is administered at selected sites domestically and internationally. Site locations are available at the ACT Test Center Locator. Typically, the sites are K–12 public, parochial, and private schools, as well as postsecondary institutions. To become a test center for domestic or international administrations, prospective sites must complete an establishment request, which is evaluated by ACT test administration staff, and then complete an establishment form. Each test center must undergo renewal annually.

### *4.1.4 Policies and Procedures*

#### Administration Manuals

For both paper and online administrations, ACT provides test centers with a variety of documentation to support standardized administration of the test. The administration manuals provide detailed directions for selecting staff, maintaining test security, and administering tests in a standardized manner. The manuals cover topics such as

- policies and procedures to follow before, during, and after testing;

- staffing levels and the responsibilities of test center staff;
- prohibited behaviors;
- handling and documenting testing irregularities;
- documentation to be submitted to ACT after testing; and
- procedures for returning test materials to ACT.

Every test center staff member must read the documentation before test day and adhere to standardized procedures.

### Staffing

The test coordinator is responsible for providing both the facilities and test center staff (room supervisors and proctors). In the event a National or International center must cancel a test date, the test coordinator must notify ACT test administration staff immediately so ACT can secure alternate facilities and staff.

### Training Staff

For standardized testing to occur successfully, all staff members must understand ACT policies and procedures and their own responsibilities for implementing them. It is critical that the same procedures be followed at every test center. The test coordinator is responsible for providing test center staff with the proper manuals and training before test day.

All staff, both new and experienced, must attend a training session conducted by the test coordinator before test day to discuss policy, procedural, and logistical issues and ensure that everyone has a common understanding of what will take place on test day.

A staff briefing session is required each test day morning, even with experienced staff. This is the time to ensure that all staff are present and make any necessary adjustments to staff assignments. The test coordinator should make sure that testing staff understand their responsibilities and should answer questions in a group setting so everyone has the same information at the same time.

All staff are required to administer and supervise the ACT in a nondiscriminatory manner and in accordance with all applicable laws, including the Americans with Disabilities Act.

## 4.2 Test Security

To ensure the validity of ACT test score interpretations, examinees, any individuals who have a role in administering the tests, and those who are otherwise involved in facilitating the testing process must strictly observe ACT's standardized testing policies and procedures. These include the test security principles set forth in ACT's administration manuals, which may be supplemented by ACT from time to time with additional communications to examinees and testing staff.

ACT's test security requirements are designed to ensure that examinees have equal opportunities to demonstrate their academic achievement and skills, that examinees who do their own work are not disadvantaged by examinees who do not, and that scores reported for each examinee have valid interpretations. Strict observation of the test security requirements is necessary to safeguard validity.

Testing staff must protect the confidentiality of ACT test items and responses. Testing staff should be aware of their responsibilities and be competent to undertake their roles, including understanding ACT's test administration policies and procedures and acknowledging and avoiding conflicts of interest in their roles as test administrators for the ACT.

Testing staff must be alert to activities that can compromise the fairness of the test and the validity of score interpretations. Such activities include, but are not limited to, cheating and questionable test-taking behavior (such as copying answers or using prohibited electronic devices during testing), accessing questions prior to the test, taking photos or making copies of test questions or test materials, posting test questions on the internet, and test proctor or test administrator misconduct (such as providing questions or answers to examinees or permitting them to engage in prohibited conduct during testing).

In addition to establishing these security-related administration protocols, ACT engages in additional test security practices designed to protect ACT test content and the validity of score interpretations. These practices include (a) establishing a reporting hotline to ACT through which individuals can anonymously report information about misconduct on an ACT test, (b) conducting data forensics to detect and respond to possible misconduct, and (c) monitoring the web to detect testing misconduct, possible unauthorized disclosure of secure ACT test content, and any other activity that might compromise the security of the ACT test or the validity of score interpretations.

## 4.3 Information Security

ACT is committed to ensuring the security, confidentiality, integrity, and availability of both its own and its customers' information, which is stored or used in support of ACT's mission to help its customers achieve education and workplace success. ACT is an information-driven company. The critical assets of the organization, including its intellectual property, are primarily information-based, and therefore ACT has a responsibility to its customers and other stakeholders to protect the information assets it owns, controls, and stores. ACT's information security manual defines ACT's governing principles for protecting information assets in all forms, whether stored electronically, on paper, or in any other form.

ACT follows the National Institute of Standards and Technology (NIST) Cybersecurity Framework (CSF), as well as the Center for Internet Security (CIS) control standards. ACT has clear rules on personally identifiable information (PII) that comply with industry standards and require encryption in source systems. ACT stores all data securely and commits to reporting and releasing data in accordance with the ACT Privacy Policy.

ACT performs comprehensive annual internal and external penetration testing. Deficiencies are identified, given remediation plans, and retested after remediation. ACT also performs daily vulnerability scans to maintain system security and privacy.

ACT has a comprehensive Security Incident Response Plan (SIRP) in place. This plan brings needed resources together in an organized manner to deal with an incident classified as an adverse event related to the safety and security of ACT networks, computer systems, and data resources. The plan ensures appropriate investigation and root cause analysis are completed if an incident takes place.

ACT requires all employees to enter confidentiality agreements and annually review ACT's information security policy. This policy appropriately classifies data and states the requirements for protecting data. Moreover, each vendor and contractor with access to ACT tests and test items is required to enter a confidentiality agreement with ACT and agree to our information security policy.

ACT is positioned to support its customers in meeting their FERPA requirements and is compliant with COPPA. ACT is not considered a covered entity subject to HIPAA.

## 4.4 Test Administration and Accessibility Levels of Support

The accessibility supports permitted during testing are designed to remove barriers to examinee access to the test yet still honor the constructs being measured. It is important to abide by all outlined requirements for administering these supports. Types of accessibility supports for the ACT include the following:

- universal supports
- designated supports
- English learner supports
- accommodations

### *4.4.1 Universal Supports*

Universal supports are available to all students and do not require ACT authorization. Examples of these embedded supports include, but are not limited to, the following:

- test booklet used as scratch paper (paper testing only)
- calculator for the mathematics section
- examinees allowed to ask for clarification of verbal instructions
- examinees allowed to ask for general administration directions to be repeated

- browser zoom/magnification (online testing only)
- “mark an item for review” function (online testing only)

### *4.4.2 Designated Supports*

Designated supports, also known as local arrangements, are available to any examinee for whom a need has been identified but whose underlying condition may not rise to the level of a disability. These supports do not require ACT approval, but most require advance planning to deliver. The test coordinator has the authority to provide the designated supports listed in the Accessibility Supports Guide for the ACT for any examinee without ACT approval so long as test security is not compromised and testing environment requirements are met. Generally, most examinees are tested with their peers; however, under certain circumstances, a change in testing conditions may be necessary for an examinee. Often, these are adjustments to the testing environment.

Examples of designated supports include, but are not limited to, the following:

- special or preferential seating or grouping
- translated test directions in limited languages
- use of approved word-to-word bilingual dictionaries

### *4.4.3 English Learner Supports*

English learner (EL) supports are available only for examinees in U.S. schools who are not proficient in English. EL supports should be identified by the educators responsible for selecting the supports needed by the student to access curriculum, instruction, and assessments because of limited English proficiency. With the exception of extra time, EL supports function as designated supports and do not require ACT approval. One-and-a-half time will be available only for multilingual students designated as having limited English proficiency and must be authorized for use by ACT prior to being used. An examinee’s English proficiency changes over time, so EL supports expire and must be reauthorized after the expiration date noted on the decision notification. Current English proficiency is measured by an English language proficiency assessment taken within the previous 2 years in the four language domains of reading, writing, speaking, and listening.

EL supports are limited to the following:

- ACT-authorized word-to-word bilingual dictionary or glossary
- translated written test directions, provided by ACT
- one-and-one-half time (requires ACT approval)
- small group testing

### 4.4.4 Accommodations

Allowed accommodations are available to users who have a documented disability. ACT requires examinees who use accommodations to have a formally documented need for, as well as relevant knowledge of and familiarity with, these supports. Accommodations must be requested and authorized by ACT in advance according to the ACT testing procedures. Appropriate documentation of the accommodation need must be provided by the examinee or by a local governing educational authority before testing.

Accommodations are available only for examinees with disabilities as documented in an IEP, 504 Plan, or another accommodations/supports plan.

Accommodations are intended to reduce or eliminate the effects of an examinee's disability; however, accommodations should never reduce learning expectations by reducing the scope, complexity, or rigor of an assessment. Accommodations provided on the ACT must be generally consistent with those provided for instruction and assessment in the educational environment. There are some accommodations that may be used in the educational environment that are not allowed for the ACT because they affect the validity of the assessment results (see 4.4.5, Modifications). There may be consequences for using prohibited or unauthorized accommodations during the ACT.

To the extent possible, accommodations should adhere to the following principles:

- Accommodations enable examinees to participate more fully and fairly in instruction and to demonstrate their knowledge and skills on the ACT.
- Accommodations are based on an examinee's need rather than on the category of an examinee's disability.
- Accommodations are based on a documented need in the instructional and assessment setting and should not be provided to give the examinee an enhancement that could be viewed as an unfair advantage or used to obtain a desired score.
- Accommodations for an examinee with disabilities are described and documented in the examinee's appropriate educational plan.
- Accommodations become part of the examinee's program of daily instruction as soon as possible after completion and approval of the educational plan.
- Accommodations are not introduced for the first time during the ACT test.
- Accommodations are monitored for effectiveness during daily instruction.

Examples of accommodations include, but are not limited to, the following:

- timing or scheduling supports (e.g., extra testing time, breaks as needed)
- presentation supports (e.g., human reading a reader's script aloud, embedded text-to-speech, compatibility with a refreshable braille device, video ASL for item content)
- response supports (e.g., scribe to record responses, computer for constructed-response items, speech-to-text software for the writing test)
- sign language interpreter for verbal instructions
- alternate formats (e.g., braille, large print)

### *4.4.5 Modifications*

Modifications are supports that are sometimes used during instruction to aid learning but, when used in a testing situation, may provide assistance in a manner that alters what the test measures. Thus, these modifications prevent the student who uses them from accessing the measured construct in the same way as examinees taking unmodified assessments. Because modifications alter the construct being tested, scores from modified assessments cannot be compared to scores from unmodified assessments. Modifications are not available for the ACT test.

For additional information on accessibility supports for the ACT, please refer to the Accessibility Supports Guide for the ACT.

# Chapter 5
# Scoring and Reporting

## 5.1 Overview

The ACT® test is composed of three multiple-choice test sections—English, mathematics, and reading—plus an optional multiple-choice science section and an optional writing test. Score reports are provided to individual students, their high schools, and colleges and other authorized recipients (e.g., scholarship agencies) of each student's choice. The contents of the student, high school, and college score reports are slightly different because they serve different purposes. The reports all contain scores indicating students' performance on each test section and detailed information about students' performance on specific areas within each section. They also provide additional information to make it easier to interpret scores and to help with college and career planning.

Chapter 2 introduced ACT scores and indicators. This chapter provides more detailed information about these scores and indicators, the writing test scoring process, and the information presented in the score reports to support college and career planning.

## 5.2 Test Section Scores, Composite Scores, STEM and ELA Scores, and Superscores

The ACT student, high school, and college reports summarize students' performance on the test sections. Results include 1–36 scale scores for each section, combined scores (the Composite score; science, technology, engineering, and mathematics [STEM] score; and English language arts [ELA] score), and, when applicable, superscores derived from a student's highest section scores across valid ACT administrations. Figures 5.1 and 5.2 show the portions of the student and high school score reports where these scores are reported. The reports also include standard errors of measurement (*SEM*s), ACT® College Readiness Benchmarks, and national (U.S.) and state ranks to aid in score interpretation.

**Figure 5.1.** Overall Score and Percentile Rank on a Sample Interactive Score Report on MyACT

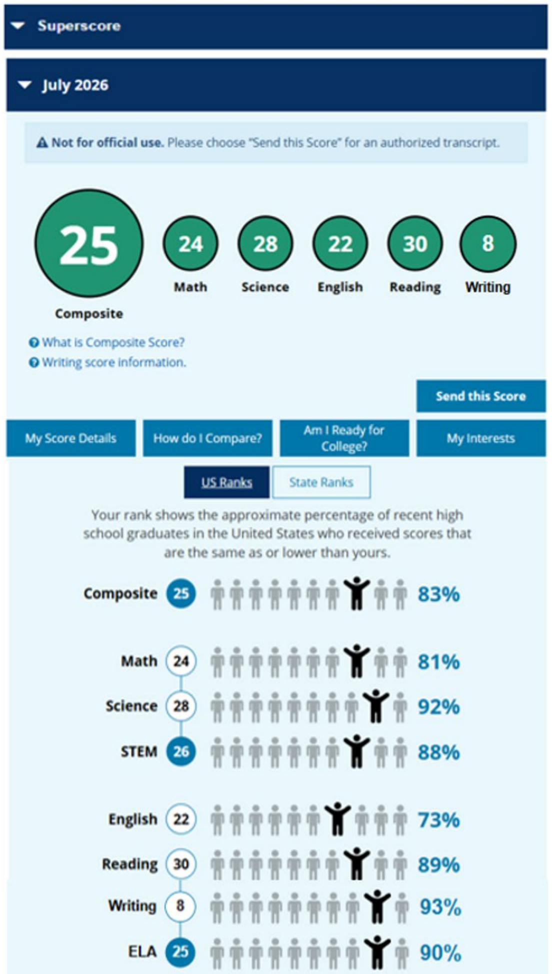
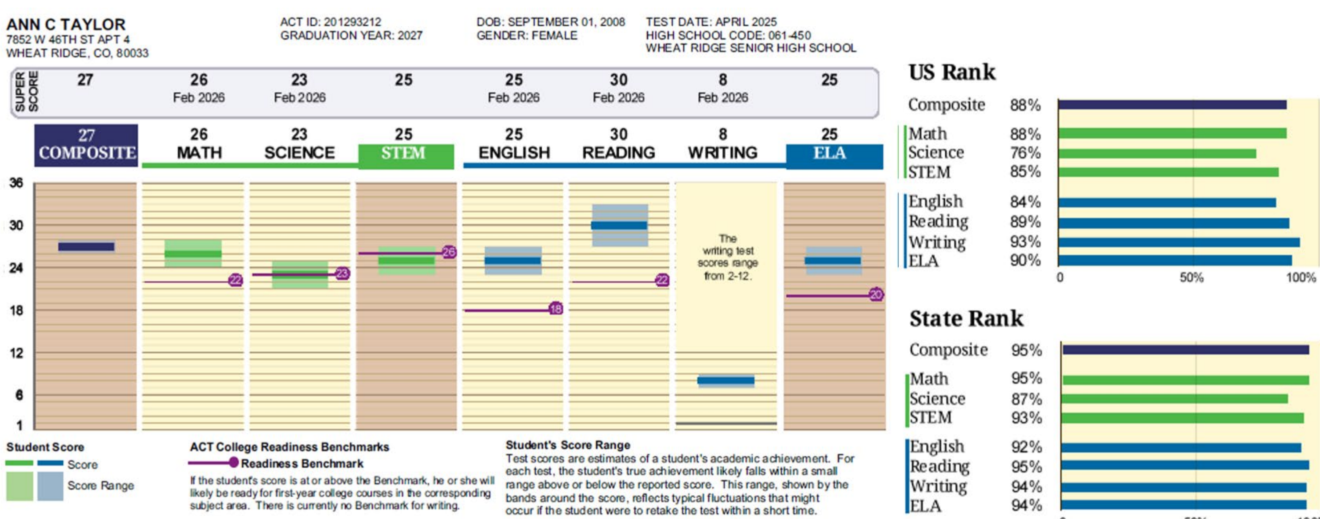



**Figure 5.2.** Overall Score and Percentile Rank on a Sample ACT High School Score Report



### 5.2.1 Test Section Scores

#### Multiple-Choice Tests

Test section scores are reported for the three multiple-choice tests (plus science if the student takes it). For each test, the raw score is the number of test questions answered correctly. Raw scores are converted to scale scores through equating procedures to ensure that scores reported across test forms have consistent meaning. Scale scores range from 1 to 36 for each of the multiple-choice tests. Procedures for obtaining the 1–36 scale scores are described in Chapter 6.

#### Writing Test Scores

If a student took the writing test, the subject-level writing test score is reported on a 2–12 scale; it is the rounded average of the four domain scores that are described in Section 5.3.2. Detailed descriptions of the domains are in Chapter 3, and the procedures used to determine the scores are described in the next section (5.2.2). Writing scores are converted to a 1–36 scale only for the purpose of calculating the ELA score; the 1–36 writing scores are not reported.

### 5.2.2 Performance Scoring for the Writing Test

All writing tests, whether administered online or on paper, are scored on online platforms. Most essays administered online are scored by one trained rater and CRASE®, ACT's automated essay scoring engine, or by CRASE alone (where eligible and when the assigned score meets an established confidence threshold). All other essays are scored independently by two trained raters. If any essay receives two scores that disagree by more than 1 point in any domain, a third rater evaluates the essay and resolves the discrepancy.

Various performance scoring processes and procedures are used for scoring the ACT writing test, such as range-finding, rater training and qualification, scoring, and rater monitoring. The scoring team responsible for these tasks is composed of raters and a scoring leadership team that includes scoring supervisors, scoring directors, content specialists, and assessment program managers. Team members' roles and responsibilities are as follows:

- Raters complete a rigorous training course and must pass a qualifying test to participate in live scoring. All raters must have, at minimum, a 4-year degree from an accredited institution of higher education. Candidates with high school English teaching experience are preferred.

- Scoring supervisors are experienced raters. Each supervisor is responsible for a team of raters. Supervisors monitor the accuracy of raters, provide feedback to raters, and resolve discrepant scores.

- Scoring directors are performance scoring professionals. Directors are responsible for the overall management of scoring work, ensuring that scores are delivered on time and meet or exceed established quality parameters.

- Content specialists and program managers form a cross-functional team of assessment development, performance scoring, and education professionals with specific expertise and credentials in English language arts. Content specialists and program managers are responsible for range-finding, training development, and ongoing calibration.

### Rater Training and Qualification

An initial range-finding process is the basis for developing scoring criteria validation. A panel of assessment and content experts meets to review a sample of student responses and ensure that content-specific criteria for the writing task accurately reflect and encompass the full range of student responses. This endeavor also establishes the foundation for subsequently developing effective rater training materials. After reviewing student responses, teams of assessment and scoring experts produce a pool of consensus-scored exemplars. At the conclusion of the process, responses are selected for specific instructional purposes and incorporated into carefully designed training sets. Developing these sets is the beginning of ACT's rigorous training program.

One item or prompt is designated for "baseline" training, which is delivered via self-paced online modules. The training modules introduce rater candidates to the rubric through the baseline training sets, beginning with a fully developed anchor set that includes exemplar responses at each score point and demonstrates a range of typical approaches to the task. After becoming familiar with anchor responses, candidates then apply scores to multiple practice sets. Practice sets include a variety of responses; some are clearly aligned with particular score points and anchor responses, while others require more detailed analysis to identify appropriate scores. A detailed annotation accompanies each baseline anchor and practice response. The annotation explains how the rubric (Table 5.1) was used to determine the score and illustrates the scoring process using citations from the exemplar response or references to appropriate anchor responses.

At the end of the training program, candidates are required to pass a qualifying test by perfectly matching a predetermined number of scores for at least two qualifying sets. Candidates who do not meet the qualifying standard are released from the scoring project.

All raters must participate in baseline training and pass the qualification test, which is administered at least once annually. After qualifying, raters are introduced to additional writing prompts via prompt-specific anchor and practice sets, but raters do not need to requalify until the following year. The pool of raters is typically a diverse group in terms of age, ethnicity, and gender, although the placement and retention of raters is based on their qualifications and the quality and accuracy of their scoring.

**Table 5.1.** Writing Test Analytic Scoring Rubric

| Score point | Ideas & Analysis | Development & Support | Organization | Language Use & Conventions |
|---|---|---|---|---|
| **Score 6: Responses at this score point demonstrate effective skill in writing an argumentative essay.** | The writer generates an argument that critically engages with multiple perspectives on the given issue. The argument's thesis reflects nuance and precision in thought and purpose. The argument establishes and employs an insightful context for analysis of the issue and its perspectives. The analysis examines implications, complexities and tensions, and/or underlying values and assumptions. | Development of ideas and support for claims deepen insight and broaden context. An integrated line of skillful reasoning and illustration effectively conveys the significance of the argument. Qualifications and complications enrich and bolster ideas and analysis. | The response exhibits a skillful organizational strategy. The response is unified by a controlling idea or purpose, and a logical progression of ideas increases the effectiveness of the writer's argument. Transitions between and within paragraphs strengthen the relationships among ideas. | The use of language enhances the argument. Word choice is skillful and precise. Sentence structures are consistently varied and clear. Stylistic and register choices, including voice and tone, are strategic and effective. While a few minor errors in grammar, usage, and mechanics may be present, they do not impede understanding. |
| **Score 5: Responses at this score point demonstrate well-developed skill in writing an argumentative essay.** | The writer generates an argument that productively engages with multiple perspectives on the given issue. The argument's thesis reflects precision in thought and purpose. The argument establishes and employs a thoughtful context for analysis of the issue and its perspectives. The analysis addresses implications, complexities and tensions, and/or underlying values and assumptions. | Development of ideas and support for claims deepen understanding. A mostly integrated line of purposeful reasoning and illustration capably conveys the significance of the argument. Qualifications and complications enrich ideas and analysis. | The response exhibits a productive organizational strategy. The response is mostly unified by a controlling idea or purpose, and a logical sequencing of ideas contributes to the effectiveness of the argument. Transitions between and within paragraphs consistently clarify the relationships among ideas. | The use of language works in service of the argument. Word choice is precise. Sentence structures are clear and varied often. Stylistic and register choices, including voice and tone, are purposeful and productive. While minor errors in grammar, usage, and mechanics may be present, they do not impede understanding. |

| Score point | Ideas & Analysis | Development & Support | Organization | Language Use & Conventions |
|---|---|---|---|---|
| **Score 4: Responses at this score point demonstrate adequate skill in writing an argumentative essay.** | The writer generates an argument that engages with multiple perspectives on the given issue. The argument's thesis reflects clarity in thought and purpose. The argument establishes and employs a relevant context for analysis of the issue and its perspectives. The analysis recognizes implications, complexities and tensions, and/or underlying values and assumptions. | Development of ideas and support for claims clarify meaning and purpose. Lines of clear reasoning and illustration adequately convey the significance of the argument. Qualifications and complications extend ideas and analysis. | The response exhibits a clear organizational strategy. The overall shape of the response reflects an emergent controlling idea or purpose. Ideas are logically grouped and sequenced. Transitions between and within paragraphs clarify the relationships among ideas. | The use of language conveys the argument with clarity. Word choice is adequate and sometimes precise. Sentence structures are clear and demonstrate some variety. Stylistic and register choices, including voice and tone, are appropriate for the rhetorical purpose. While errors in grammar, usage, and mechanics are present, they rarely impede understanding. |
| **Score 3: Responses at this score point demonstrate some developing skill in writing an argumentative essay.** | The writer generates an argument that responds to multiple perspectives on the given issue. The argument's thesis reflects some clarity in thought and purpose. The argument establishes a limited or tangential context for analysis of the issue and its perspectives. Analysis is simplistic or somewhat unclear. | Development of ideas and support for claims are mostly relevant but are overly general or simplistic. Reasoning and illustration largely clarify the argument but may be somewhat repetitious or imprecise. | The response exhibits a basic organizational structure. The response largely coheres, with most ideas logically grouped. Transitions between and within paragraphs sometimes clarify the relationships among ideas. | The use of language is basic and only somewhat clear. Word choice is general and occasionally imprecise. Sentence structures are usually clear but show little variety. Stylistic and register choices, including voice and tone, are not always appropriate for the rhetorical purpose. Distracting errors in grammar, usage, and mechanics may be present, but they generally do not impede understanding. |

| Score point | Ideas & Analysis | Development & Support | Organization | Language Use & Conventions |
|---|---|---|---|---|
| **Score 2: Responses at this score point demonstrate weak or inconsistent skill in writing an argumentative essay.** | The writer generates an argument that weakly responds to multiple perspectives on the given issue. The argument's thesis, if evident, reflects little clarity in thought and purpose. Attempts at analysis are incomplete, are largely irrelevant, or consist primarily of restatement of the issue and its perspectives. | Development of ideas and support for claims are weak, confused, or disjointed. Reasoning and illustration are inadequate, illogical, or circular, and fail to fully clarify the argument. | The response exhibits a rudimentary organizational structure. Grouping of ideas is inconsistent and often unclear. Transitions between and within paragraphs are misleading or poorly formed. | The use of language is inconsistent and often unclear. Word choice is rudimentary and frequently imprecise. Sentence structures are sometimes unclear. Stylistic and register choices, including voice and tone, are inconsistent and not always appropriate for the rhetorical purpose. Distracting errors in grammar, usage, and mechanics are present, and they sometimes impede understanding. |
| **Score 1: Responses at this score point demonstrate little or no skill in writing an argumentative essay.** | The writer fails to generate an argument that responds intelligibly to the task. The writer's intentions are difficult to discern. Attempts at analysis are unclear or irrelevant. | Ideas lack development, and claims lack support. Reasoning and illustration are unclear, incoherent, or largely absent. | The response does not exhibit an organizational structure. There is little grouping of ideas. When present, transitional devices fail to connect ideas. | The use of language fails to demonstrate skill in responding to the task. Word choice is imprecise and often difficult to comprehend. Sentence structures are often unclear. Stylistic and register choices are difficult to identify. Errors in grammar, usage, and mechanics are pervasive and often impede understanding. |

## Managing Rater Quality

Training and qualification provide initial quality assurance for all raters, but quality monitoring activities continue throughout the performance scoring process. ACT employs several quality assurance processes that establish and maintain consistent calibration and ensure that every response, whether scored on the first day or the last, is given the most appropriate score. ACT's standard quality assurance practices include the following:

- **Reliability Scoring.** Determining how often and how closely raters agree with one another ensures consistent application of the rubric. Most ACT writing responses are reviewed and independently scored by at least two qualified raters to generate inter-rater reliability data. In cases where scores are nonadjacent, a response is automatically rerouted for a third review by a scoring supervisor or director, and the discrepancy is appropriately resolved. Because of the rigorous training and qualification requirements, nonadjacency rates routinely amount to less than 4% of the overall response population.

- **Validity.** Rater accuracy is measured by the rate of agreement with prescored validity responses. Validity responses are selected and scored by scoring leaders and then embedded in the workflow so that raters cannot see the applied scores until after they have scored the response. Approved scores are then revealed to reinforce accurate scoring or redirect rater alignment to the rubric. A rater whose performance falls below established quality thresholds is excluded from scoring and is subject to retraining activities, including receiving supervisor feedback and taking calibration tests. Raters who fail to demonstrate improved accuracy may be released from the project and their work reset and rescored.

- **Backreading.** The backreading process enables scoring leaders to review raters' work and provide customized feedback based on specific scoring examples. This process also allows for new scores to be applied where necessary. This is an important part of the quality assurance process, and all raters are subject to daily backreading.

- **Calibration.** Scoring leaders create calibration and retraining exercises to reinforce range-finding standards, communicate scoring decisions, correct scoring issues and trends, or transition to a new item. Calibration sets can be targeted to individual raters, groups of raters, or the entire scoring team. General and targeted calibration exercises are administered regularly throughout the performance scoring process.

- **Quality Reporting.** ACT uses a set of dynamic, on-demand quality reports to monitor scoring quality. Up-to-date rater performance statistics help quickly identify and diagnose scoring issues at the group or individual rater level. Scoring leaders respond to live scoring conditions on an ongoing basis by reviewing inter-rater reliability, validity agreement, frequency distribution, scoring rate, backreading agreement, and other important quality metrics. These data insights allow scoring leaders to determine how to address any issues through ongoing training and intervention. Table 5.2 provides a sample of some of the available reports.

**Table 5.2.** Sample of Quality Reports

| Report name | Description |
|---|---|
| Inter-Rater Reliability Summary | Group-level summary of inter-rater reliability statistics. These statistics are calculated by measuring the agreement between the first and second scores, averaged for the group. |
| Frequency Distribution Report | Item-level summary of score-point distribution percentages on both a daily and a cumulative basis. |
| Daily/Cumulative Validity Summary | Summary of agreement for validity reads of a given task on both a daily and a cumulative basis. |
| Completion Report | Breakdown of the number of responses scored and the number of responses in each stage of scoring (first score, second score, resolution). |
| Daily Scoring Management | Summary of item-level validity and inter-rater reliability on a daily and cumulative basis; this report also shows the number of resolutions required and completed, as well as item-level frequency distribution. |

### CRASE Scoring

For the online ACT writing test, ACT uses both trained human raters and its constructed response automated scoring engine (CRASE). More information about CRASE is available on the CRASE web page.

Most online responses can be scored by one human rater and CRASE, or by CRASE alone. The CRASE scoring models for ACT writing were created using around 9,000 human-scored essays across a variety of prompts. The models produced scores that agreed with human rater scores at rates that matched or exceeded the agreement rates between scores produced by independent human raters. Details about the training and validation process, including information about model performance across subgroups, can be found in the document *CRASE5 for ACT Writing Technical Report*.

ACT also uses automated scoring technologies that assign a single automated score set to online responses while maintaining the reliability and validity of performance scoring results. The CRASE Confidence Score process relies on the CRASE platform, which generates confidence ratings for essays. Essays that closely match established model parameters receive the highest ratings and are sent directly to score calculation and reporting without further intervention. Essays with lower confidence ratings, along with a random sample of essays with high confidence ratings, are reviewed by human raters following the standard scoring process.

### *5.2.3 Composite, STEM, and ELA Scores*

The ACT Composite score represents a student's overall performance on the test. Beginning with the enhanced ACT, science is reported separately and is not included in the Composite score calculation, which is the average of the scale scores for English, mathematics, and reading, rounded to the nearest whole number. As a result, the Composite score can be reported for all examinees, regardless of whether they take the optional science test.

The STEM score represents a student's overall performance on the mathematics and science sections. It is the rounded average of the mathematics and science scale scores, so only students who receive both mathematics and science scores receive a STEM score.

The ELA score represents a student's overall performance across the English, reading, and writing sections. It is the rounded average of the English, reading, and 1–36 writing scale scores, so only students who take the writing test along with the ACT test receive an ELA score. To calculate ELA scores, ACT converts the sum of the writing domain scores to a 1–36 scale. The 1–36 writing scale score is used only for calculating the ELA score and is not reported independently. Procedures for obtaining the 1–36 writing scale score are described in Chapter 6.

The Composite, STEM, and ELA scores all range from 1 to 36. Because the component section scale scores are equated, the derived Composite, STEM, and ELA scores have consistent interpretations across ACT test forms.

### *5.2.4 ACT Superscores*

The ACT Superscore is calculated like the ACT Composite score but is based on the student's highest ACT section scores across multiple test administrations. Beginning with the enhanced ACT, the ACT Superscore is based on the highest English, mathematics, and reading scores from eligible administrations; these scores are averaged, and the result is rounded to the nearest whole number. Science is not included, although a science score continues to be reported separately when a student takes the science test.

The same general approach is used to calculate STEM and ELA Superscores when the required component scores are available (mathematics and science for STEM; English, reading, and writing for ELA).

Superscores were first provided to students during the 2020–2021 academic year, and they count as official ACT scores for reporting to colleges and universities. Research on the validity of ACT Superscores is provided in Chapter 7.

### *5.2.5 Interpretation of ACT Test Scores*

The ACT score reports present additional information to help students and educators interpret scores. This includes standard errors of measurement (*SEM*s), the ACT College Readiness Benchmarks, and the national and state score ranks.

### *SEM* and Score Ranges

The score report contains information about the measurement precision of the Composite, test section, STEM, and ELA scores. The *SEM* reflects imprecision in test scores related to the fact that students would not necessarily earn the same scores if they took the ACT repeatedly. The *SEM*s are about 1 point for the writing and Composite scores; about 2 points for the English, mathematics, and science sections, the STEM score, and the ELA score; and about 3 points for the reading section. Students' reported scores are accompanied by score ranges based on the *SEM*s and are graphically represented by shaded areas around their scores. These ranges are intended to help users avoid overinterpreting small differences between scores. Detailed information about measurement precision is given in Chapter 6.

### ACT College Readiness Benchmarks

The ACT College Readiness Benchmarks are scores that represent the level of achievement associated with at least a 50% chance of earning a B or higher or about a 75% chance of earning a C or higher in corresponding credit-bearing first-year college courses. Using data collected from hundreds of colleges across the United States, ACT has established Benchmarks for the ACT section test scores, the ACT STEM score, and the ACT ELA score.

Students' readiness for first-year college courses corresponding to the English, mathematics, reading, and science tests, as well as to the STEM and ELA scores, can be assessed by comparing students' scores with these Benchmarks. The STEM Benchmark is the minimum STEM score associated with readiness for first-year college courses typically taken by STEM majors, and the ELA Benchmark is the minimum ELA score associated with readiness for first-year college ELA courses.

Additional resources are available to facilitate interpretation of ACT scores. The ACT® College and Career Readiness Standards are research-based sets of statements that describe the essential skills and knowledge students need for college and career readiness. They are intended to help students, parents, and educators understand the meaning of ACT test scores by relating scores to the types of skills needed for success in high school and beyond. The Standards serve as a direct link between what students have learned and what they are ready to learn next. They are informed by ACT's ongoing research program, including the ACT® National Curriculum Survey®, which gathers evidence from K–12 and postsecondary educators about instructional practices, expectations, and readiness requirements. ACT also provides Standards documents and curriculum review worksheets for English, mathematics, reading, science, and writing on the ACT College and Career Readiness Standards web page to support interpretation and instructional planning. For additional information about interpreting ACT scores, the College Readiness Benchmarks, and the College and Career Readiness Standards, see Sections 5.5 and 5.6.

### Score Norms

National (U.S.) and state ranks support norm-referenced interpretations by indicating how an examinee's scores compare with those of other examinees in the nation and in their state. A rank indicates the percentage of examinees in a specified reference group who obtained scores at or below a given score. The ranks are based on observed score distributions from recent cohorts and are provided to support score interpretation and comparisons among examinees.

ACT national and state ranks are based on the most recent scores from a rolling 3-year cohort of ACT-tested high school graduates who tested in Grades 10, 11, or 12; national ranks are derived from this national reference group, and state ranks are based on the subset of examinees within a given state. Because the reference group includes examinees tested before Grade 12, these ranks are not intended to represent the performance of a specific graduating class; rather, they summarize the performance of a recent cohort of ACT-tested students.

Norms also support comparisons across test sections (English, mathematics, reading, science, and writing). Scale scores are defined independently for each test section and cannot be directly compared across section; for example, a scale score of 22 on two different test sections does not indicate equivalent performance. In contrast, ranks express performance relative to common reference groups, allowing meaningful comparisons of an examinee's relative standing across test sections.

The most recent national ranks are available online at ACT Test Scores: National Ranks.

### *5.2.6 Summary Statistics, Effective Weights, and Correlations*

Operational test data from eight test forms administered from June 2025 to April 2026 were analyzed to obtain the descriptive statistics reported in this chapter. The data included large national samples. This section presents the summary statistics and correlations among the ACT section scores and the Composite, STEM, and ELA scores. Effective weights are also reported for each component of the Composite, STEM, and ELA scores. For the enhanced ACT, the Composite score is based on English, mathematics, and reading; science is offered as an optional test and is reported separately when taken. STEM scores are reported when examinees have eligible mathematics and science scores. ELA scores are reported when examinees have eligible English, reading, and writing scores.

#### Score Distribution Summary Statistics

The summary statistics of the ACT test scores on eight forms administered from June 2025 to April 2026 are presented in Table 5.3. Summary statistics for STEM are based on examinees with eligible mathematics and science scores, and summary statistics for ELA are based on examinees with eligible English, reading, and writing scores.

Table 5.4 provides corresponding statistics for the best test section score and ACT Superscore distributions calculated using data from the same students. For students who tested only once, their single test attempts were used for superscoring. Beginning with the enhanced ACT, the ACT Superscore is calculated using the student's highest English, mathematics, and reading scores. Since the Superscore reflects the best test section scores for each student, the Superscore means in Table 5.4 are slightly higher than the means in Table 5.3.

**Table 5.3.** Summary Statistics of the ACT Test Score Distributions

| Statistic | English | Math | Reading | Science | Writing | Comp. | STEM | ELA |
|---|---|---|---|---|---|---|---|---|
| **Mean** | 22.64 | 22.70 | 24.37 | 24.06 | 7.08 | 23.24 | 23.94 | 20.99 |
| ***SD*** | 6.69 | 5.88 | 6.75 | 5.57 | 1.63 | 5.79 | 5.42 | 5.31 |
| **Skewness** | 0.09 | 0.29 | −0.18 | 0.07 | −0.05 | 0.06 | 0.15 | 0.02 |
| **Kurtosis** | 2.33 | 2.11 | 2.28 | 2.59 | 3.10 | 2.19 | 2.24 | 2.51 |

**Table 5.4.** Summary Statistics of the Best ACT Test Section Score and ACT Superscore Distributions

| Statistic | English | Math | Reading | Science | Writing | Comp. | STEM | ELA |
|---|---|---|---|---|---|---|---|---|
| **Mean** | 23.18 | 23.17 | 25.25 | 24.65 | 7.16 | 23.87 | 24.45 | 22.50 |
| ***SD*** | 6.70 | 5.86 | 6.62 | 5.53 | 1.60 | 5.85 | 5.45 | 5.31 |
| **Skewness** | 0.08 | 0.25 | −0.22 | 0.07 | −0.08 | 0.02 | 0.13 | −0.11 |
| **Kurtosis** | 2.30 | 2.11 | 2.29 | 2.53 | 3.13 | 2.18 | 2.22 | 2.49 |

### Effective Weights

The Composite, STEM, and ELA scores are rounded averages of test section scores. Specifically, the English, mathematics, and reading scale scores are weighted equally to calculate the Composite score (science is not included); the mathematics and science scale scores are weighted equally to calculate the STEM score; and the English, reading, and writing scale scores are weighted equally to calculate the ELA score. This means the weights used in the calculation are one third for the Composite score, one half for the STEM score, and one third for the ELA score (often referred to as "nominal" weights).

There are other ways to determine the contributions of test scores to a combined score. The effective weight, for example, is the proportion of the variability of the combined score that can be attributed to a particular test score (Wang & Stanley, 1970). Score covariances are calculated and combined to obtain effective weights. Specifically, the effective weight for a test is calculated by summing the values in the appropriate row of the covariance matrix and dividing the resulting value by the sum of all covariances among the tests using the formula

$$(effective\ weight)_x = \frac{\sum_y cov_{xy}}{\sum_x \sum_y cov_{xy}},$$

where $cov_{xy}$ is the covariance of test scores corresponding to row $x$ and column $y$ in the covariance matrix.

As an illustration, Table 5.5 shows the scale score covariances for the English, mathematics, and reading tests for one of the ACT test forms administered between June 2025 and April 2026. These values are used to calculate the effective weight of each test section for the Composite score. For example, the effective weight of the English test was computed by adding the three numbers in the first row (42.62, 25.89, and 32.93) and then dividing this number by the sum of all covariances for the three multiple-choice tests, which resulted in an effective weight of 0.36 (after rounding). The effective weights for the mathematics and reading tests were obtained in a similar fashion.

Table 5.6 shows the ranges of effective weights for the Composite, STEM, and ELA scores based on the eight test forms administered from June 2025 to April 2026. The relative sizes of the effective weights depend on the score variances and covariances among the component scores. For example, component tests with larger score variances or stronger covariances with the other component tests tend to have larger effective weights.

**Table 5.5.** Scale Score Covariance Matrix for Multiple-Choice Tests From One ACT Test Form

| Test | English | Math | Reading |
|---|---|---|---|
| English | 42.62 | 25.89 | 32.93 |
| Math | 25.89 | 31.61 | 23.36 |
| Reading | 32.93 | 23.36 | 45.26 |

**Table 5.6.** Range of Effective Weights for Composite, STEM, and ELA Scores

| Test | Composite | STEM | ELA |
|---|---|---|---|
| English | 0.35–0.36 | — | 0.36–0.40 |
| Math | 0.28–0.30 | 0.50–0.54 | — |
| Reading | 0.34–0.36 | — | 0.36–0.41 |
| Science | — | 0.46–0.50 | — |
| Writing | — | — | 0.20–0.27 |

### Correlations

Table 5.7 shows the correlations among ACT section and Composite, STEM, and ELA scores based on operational data from eight test forms administered from June 2025 to April 2026. The correlations between the writing scores and other scale scores were relatively low, which was attributable to the writing scores having a smaller range and lower reliability than the other scores. Reliability estimates for ACT scores are presented in Chapter 6.

**Table 5.7.** Correlations Among the ACT Test Scores

| Test | English | Math | Reading | Science | Composite | STEM | Writing | ELA |
|---|---|---|---|---|---|---|---|---|
| **English** | 1.00 | 0.72 | 0.77 | 0.73 | 0.93 | 0.77 | 0.49 | 0.90 |
| **Math** | — | 1.00 | 0.64 | 0.77 | 0.86 | 0.94 | 0.40 | 0.68 |
| **Reading** | — | — | 1.00 | 0.70 | 0.90 | 0.71 | 0.45 | 0.90 |
| **Science** | — | — | — | 1.00 | 0.82 | 0.94 | 0.44 | 0.75 |
| **Composite** | — | — | — | — | 1.00 | 0.89 | 0.50 | 0.93 |
| **STEM** | — | — | — | — | — | 1.00 | 0.45 | 0.77 |
| **Writing** | — | — | — | — | — | — | 1.00 | 0.71 |
| **ELA** | — | — | — | — | — | — | — | 1.00 |

## 5.3 Detailed Performance Description

As shown in Figures 5.3 and 5.4, ACT score reports include detailed results that describe students' performance on finer-grained skills and domains within each test section. This includes reporting category scores and ACT Readiness Ranges for each multiple-choice test, as well as domain scores for the ACT writing test.

**Figure 5.3.** Detailed Results on a Sample Interactive Score Report on MyACT

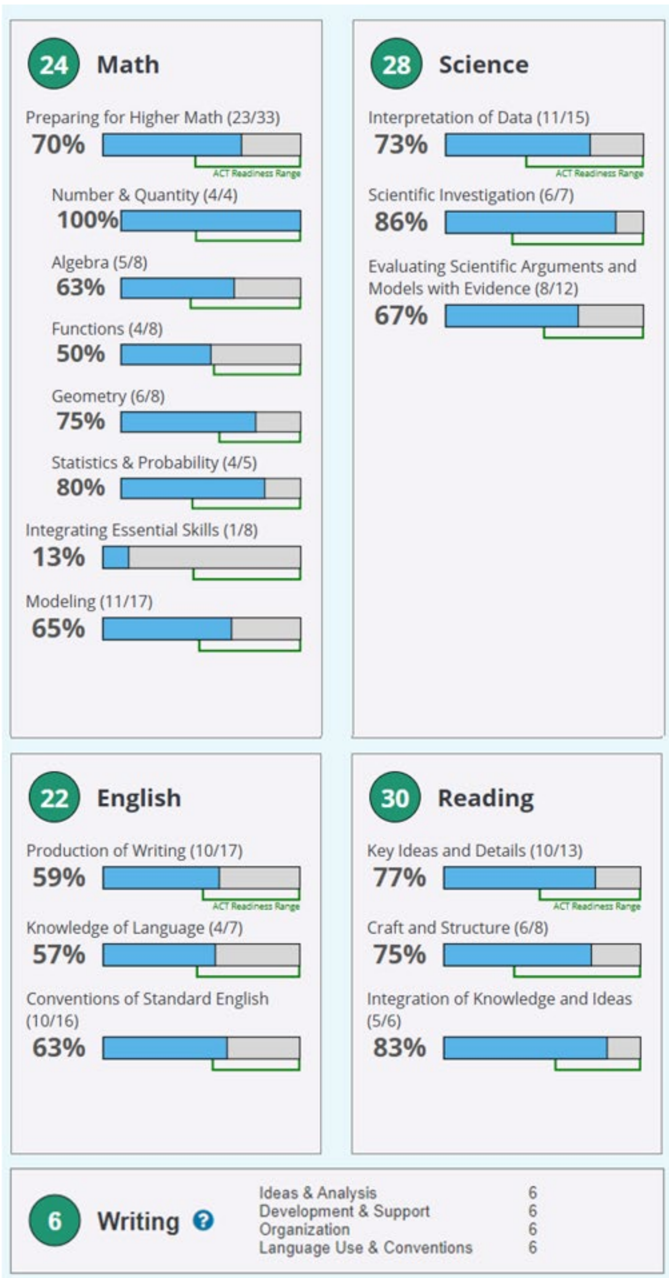
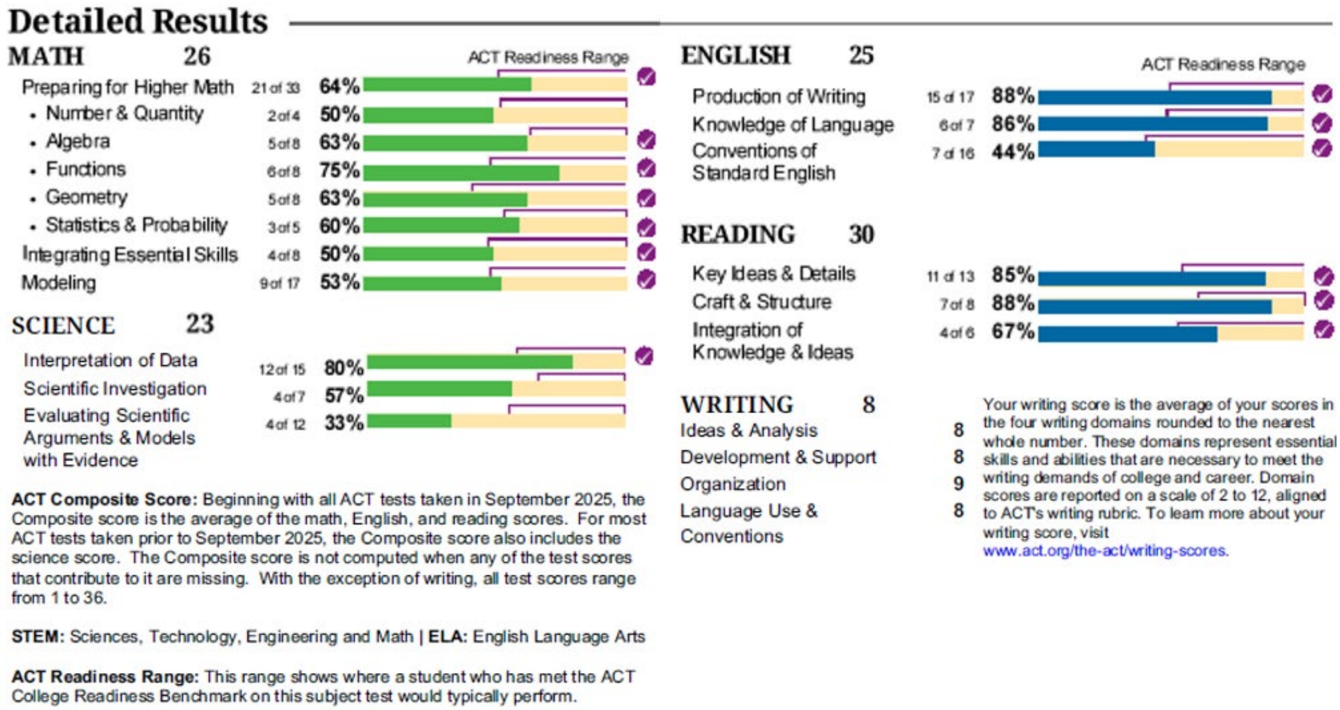


**Figure 5.4.** Detailed Results on a Sample ACT High School Score Report




### *5.3.1 Reporting Categories and ACT Readiness Ranges*

The ACT reporting categories are aligned with the ACT College and Career Readiness Standards (see Section 5.5) and other standards that target college and career readiness. Items that measure similar skills are grouped together to provide students with more detailed information about their test performance within each section. There are three reporting categories each for English, reading, and science and eight for mathematics. These reporting categories make it easier for students, parents, and educators to gain insight into students' performance by highlighting students' relative strengths and areas for improvement in each section.

For each reporting category, the total number of points possible, the total number of points a student obtained, and the percentage of points achieved are shown in the detailed results that are included in the score reports (Figures 5.3 and 5.4). In addition, for each reporting category, there is an ACT Readiness Range indicating the percentage-correct scores expected for students who scored at or above the ACT College Readiness Benchmark for that specific section. Note that the number of items for a particular reporting category can vary across different test forms. The Readiness Ranges vary accordingly, and they also account for differences in reporting category item difficulty across forms following the procedure described in Chapter 6.

Information about the development and blueprints of ACT reporting categories is in Chapter 3. Details about interpreting ACT reporting categories and ACT Readiness Ranges are in the *ACT Reporting Category Interpretation Guide* (Powers et al., 2016).

### *5.3.2 Writing Domain Scores*

In addition to the overall writing test score, scores are reported for four domains: Ideas & Analysis, Development & Support, Organization, and Language Use & Conventions. These domains reflect essential skills and abilities in writing that are required for college and career success (see Chapter 3 for detailed descriptions of the domains). The rubric used to score the writing section reflects a 1–6 scale (see Table 5.1); however, each domain receives a final score of 2–12, either calculated as the sum of two raters' scores or assigned by the automated scoring engine.

Table 5.8 presents the summary statistics of writing domain scores and the overall writing scores based on ACT national writing test forms administered from June 2025 to April 2026. Table 5.9 presents the correlations among these scores for ACT national testers.

**Table 5.8.** Summary Statistics of the ACT Writing and Writing Domain Score Distributions

| Statistic | Ideas & Analysis | Development & Support | Organization | Language Use & Conventions | Writing score |
|---|---|---|---|---|---|
| Mean | 7.04 | 6.46 | 6.92 | 7.37 | 7.08 |
| *SD* | 1.74 | 1.76 | 1.67 | 1.51 | 1.63 |
| Skewness | –0.17 | 0.05 | –0.23 | 0.04 | –0.05 |
| Kurtosis | 3.12 | 2.73 | 3.14 | 3.38 | 3.10 |

**Table 5.9.** Correlations Among the ACT Writing and Writing Domain Scores

| Score | Ideas & Analysis | Development & Support | Organization | Language Use & Conventions | Writing score |
|---|---|---|---|---|---|
| Ideas & Analysis | 1.00 | 0.92 | 0.97 | 0.94 | 0.98 |
| Development & Support | — | 1.00 | 0.93 | 0.89 | 0.93 |
| Organization | — | — | 1.00 | 0.92 | 0.97 |
| Language Use & Conventions | — | — | — | 1.00 | 0.96 |
| Writing score | — | — | — | — | 1.00 |

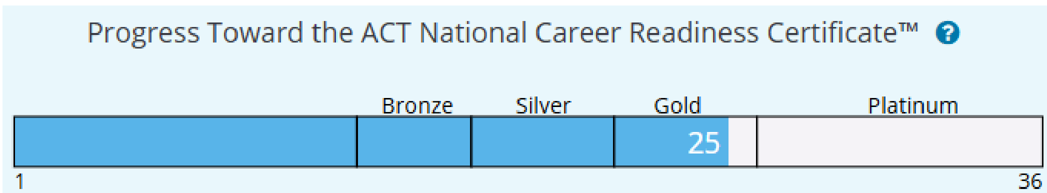
## 5.4 Progress Toward the ACT WorkKeys National Career Readiness Certificate Indicator

The Progress Toward the ACT® WorkKeys® National Career Readiness Certificate™ (NCRC®) indicator is based on students' ACT Composite scores. This indicator provides an estimate of students' most likely performance on the ACT WorkKeys NCRC, an assessment-based credential that certifies foundational workplace skills required for success across a wide range of industries and occupations.

The NCRC is awarded at four levels—Bronze, Silver, Gold, and Platinum—based on performance on three assessments: ACT WorkKeys Applied Math, Graphic Literacy, and Workplace Documents. The credential provides evidence that individuals possess essential skills valued by employers. Find more information about the NCRC at ACT WorkKeys National Career Readiness Certificate (NCRC).

Figure 5.5 shows the indicator as it appears on the ACT High School Score Report. The sample indicator provides an estimate of the NCRC level a student with a 25 ACT Composite score is likely to obtain (Gold).

**Figure 5.5.** ACT High School Report Indicator of Progress Toward the ACT WorkKeys National Career Readiness Certificate




The NCRC indicator predicts the NCRC level students are likely to achieve if they complete the WorkKeys assessments at approximately the same time as the ACT test. For the 2025–2026 testing year, ACT updated this indicator using recent data from students who completed both the ACT and all three WorkKeys assessments. The update reflects changes to the ACT Composite score calculation, which now represents the average of the English, mathematics, and reading scores.

The updated indicator is based on a sample of 338,832 students who took both the ACT and WorkKeys assessments within 1 academic year between fall 2022 and summer 2025. The average ACT Composite score of the dataset was 18.5. The sample was weighted to balance the different order in which the tests were taken. Logistic regression and classification procedures were used to identify Composite score ranges corresponding to each NCRC level, defined as the score ranges at which at least 50% of students achieved a given NCRC level or higher. This approach is consistent with earlier linking methodologies used to relate ACT scores to external performance indicators.

Based on these analyses, the ACT Composite score ranges associated with predicted NCRC levels were updated in fall 2025. The resulting classifications are as follows:

- Nonqualifier: 1–12
- Bronze: 13–15
- Silver: 16–20
- Gold: 21–26

- Platinum: 27–36

These updated ranges differ slightly from those used before fall 2025, reflecting changes in the ACT Composite score definition and the use of more recent linking data.

The NCRC indicator is not a substitute for an actual NCRC level obtained through direct assessment with WorkKeys. Because the indicator is based on probabilistic relationships between ACT scores and WorkKeys outcomes, individual results may differ from predicted outcomes. In addition, although both the ACT and WorkKeys assessments provide information about readiness, they measure related but distinct constructs and content domains.

Despite these limitations, the NCRC indicator provides useful information about students' likely level of career readiness based on their ACT performance and supports interpretation of ACT scores in relation to workplace skill development.

## 5.5 ACT College and Career Readiness Standards

This section provides information about the ACT College and Career Readiness Standards—what their purpose is, how they were developed, how they are maintained, and how to interpret them. These Standards are empirically derived descriptions of the essential skills and knowledge students are likely to demonstrate at various score levels on ACT tests. Parents, teachers, counselors, and students use the Standards to

- communicate widely shared learning goals and expectations;
- relate ACT test scores to the knowledge and skills needed for success in high school, college, and careers; and
- understand the increasing complexity of the skills needed across the score ranges in English, mathematics, reading, science, and writing.

The ACT College Readiness Benchmarks, described in Section 5.6, complement the Standards by identifying the minimum ACT scores associated with a reasonable probability of success in selected first-year credit-bearing college courses.

### *5.5.1 Description of the ACT College and Career Readiness Standards*

ACT developed the College and Career Readiness Standards in 1997 to provide a more meaningful interpretation of ACT scores. The Standards describe what students who score within particular score ranges are likely to know and be able to do in English, mathematics, reading, science, and writing. Students scoring within lower score ranges typically demonstrate foundational skills, whereas students scoring within higher score ranges generally demonstrate the ability to apply those skills in more complex situations. The Standards therefore provide a progression of skills and knowledge across score ranges within each subject area.

The Standards are organized by reporting categories, or strands, within each test to make them easier for teachers and curriculum specialists to use. Current versions of the Standards are available at ACT College and Career Readiness Standards.

For the ACT multiple-choice tests, Standards are provided for the following score ranges: 13–15, 16–19, 20–23, 24–27, 28–32, and 33–36. Students scoring in the 1–12 range are generally beginning to develop the knowledge and skills described in the 13–15 range; therefore, separate Standards are not reported for scores below 13. The Standards are cumulative: Students scoring in a given range are generally expected to demonstrate most or all of the knowledge and skills described in the lower ranges in addition to the skills associated with their own range.

ACT developed Standards for the writing test in 2005 and updated them in 2015. Writing Standards are reported by writing domain and score range: 3–4, 5–6, 7–8, and 9–10. Scores below 3 in a writing domain do not provide sufficient evidence to support meaningful generalizations about the writing skills associated with that domain. Current versions of the ACT writing Standards are available at ACT Writing College and Career Readiness Standards.

### *5.5.2 Determining the Score Ranges for the ACT College and Career Readiness Standards*

During the initial development of the Standards, normative data, college admissions criteria, and course-placement research were reviewed to establish score ranges that would meaningfully distinguish levels of student achievement. The analyses were conducted to identify score ranges that would help educators, students, and other stakeholders interpret ACT scores in terms of the knowledge and skills students were likely to demonstrate. Findings from admissions and course-placement research were used to ensure that the resulting score ranges aligned with important educational decisions and outcomes.

Based on these analyses, ACT established the score ranges 1–12, 13–15, 16–19, 20–23, 24–27, 28–32, and 33–36. These ranges were selected because they provided meaningful distinctions among levels of achievement while supporting clear interpretations of student knowledge and skills. The six reported Standards ranges (13–15 through 33–36) continue to serve as the framework for describing increasingly sophisticated levels of achievement on ACT tests.

### *5.5.3 Developing the ACT College and Career Readiness Standards*

The Standards were developed through a multistage process that combined empirical analyses of student performance with expert review of assessment content. Following the establishment of the score ranges described in Section 5.5.2, content specialists in English, mathematics, reading, and science reviewed items from multiple ACT test forms and identified skills and knowledge necessary to answer the test items in each score range.

For each score range, the content specialists examined test items answered correctly by at least 80% of examinees scoring within that range. This criterion was selected because it provided a high degree of confidence that students achieving scores within a particular range were likely to possess the knowledge and skills required to answer those items successfully.

Using item performance data and detailed reviews of test content, the content specialists identified the skills and knowledge typically measured by items associated with each score range. These observations were synthesized into score range descriptions representing increasingly sophisticated levels of achievement. Multiple test forms were reviewed to ensure that the resulting Standards reflected the broader content domains assessed by the ACT rather than the characteristics of any single test form.

Because students in higher score ranges generally answer items associated with lower score ranges correctly, the Standards were designed to be cumulative. Students scoring in higher ranges are therefore expected to demonstrate most or all of the skills and knowledge associated with preceding score ranges.

The draft Standards underwent multiple rounds of review and refinement. The preliminary Standards were reviewed against additional ACT test forms, and adjustments and refinements were made as needed. Content specialists reviewed the revisions, and the updated Standards were subsequently evaluated using additional test forms to determine whether they adequately described the knowledge and skills associated with each score range. Further refinements were made before the Standards were finalized.

To gather evidence supporting the content validity of the Standards, ACT conducted an independent national review involving nationally recognized scholars, teachers, and curriculum specialists in English, mathematics, reading, and science. The review panel included college faculty and classroom teachers with extensive and current knowledge of their disciplines and curriculum, and many participants had received national recognition for their professional accomplishments. The review process sought broad representation across genders, ethnic backgrounds, and geographic locations.

The reviewers were shown the complete set of Standards and evaluated (a) whether the Standards accurately described the knowledge and skills required to respond successfully to ACT test items within each score range and (b) whether the Standards represented a continuum of increasingly sophisticated knowledge and skills across score ranges. To support these evaluations, the reviewers were given sample ACT test items associated with each score range. Item samples were selected randomly to promote an objective review of the correspondence between the Standards and the skills measured by the assessments.

Feedback from the independent review informed refinements to the language and organization of the Standards. The resulting Standards provide a research-based framework for interpreting ACT scores in terms of the knowledge and skills students are likely to demonstrate.

### *5.5.4 The ACT College and Career Readiness Standards for Writing*

The score ranges and the Standards for the writing test were derived from the ACT writing test scoring rubric, a four-domain, 6-point descriptive scale used to determine writing test scores

(Table 5.1). See Section 5.2.2 for a detailed description of the scoring process and Section 5.3.2 for a description of how domain scores are produced.

The domains assessed by the ACT writing test correspond to key dimensions of effective writing that are taught in high school and college-level composition courses: Ideas & Analysis, Development & Support, Organization, and Language Use & Conventions. These writing domains replace the previous five strands of the writing Standards, which were derived from a holistic scoring rubric. The design of the writing test and accompanying Standards reflects the input of several independent consultants, including high school and postsecondary instructors, as well as results from the ACT National Curriculum Survey.

To determine the score ranges for the writing Standards, ACT staff considered the differences in writing ability evident in essays that scored between levels of the scoring rubric (e.g., 6.5). Based on similarities found among written responses at certain adjacent score points, ACT staff determined that five score ranges would best distinguish students' levels of writing achievement to help teachers, administrators, and others relate ACT test scores to students' skills and knowledge. Writing that receives a score of 3 or lower does not permit useful generalizations about the student's writing abilities in that domain.

### *5.5.5 Periodic Review of the ACT College and Career Readiness Standards*

ACT has conducted periodic additional reviews of its College and Career Readiness Standards. For those reviews, ACT identified three to four new test forms and analyzed the data and the corresponding test items specific to each score range. Topics were also compared to data from the ACT National Curriculum Survey (ACT, 2007a, 2009, 2013, 2016a, 20). The purpose of these reviews was to ensure that the Standards reflected (a) the most important knowledge and skills for college and career readiness, (b) what was being measured by the items in each score range, and (c) a cumulative progression of increasingly sophisticated skills and knowledge from the lowest score range to the highest score range. These reviews resulted in minor refinements intended to update and clarify the language of the Standards.

### *5.5.6 Interpreting and Using the ACT College and Career Readiness Standards*

Because new ACT test forms are developed regularly and because no one test form measures all the skills and knowledge included in any particular standard, the ACT College and Career Readiness Standards must be interpreted as knowledge and skills that most students who score within a particular score range are likely to be able to demonstrate.

ACT tests include items measuring areas of knowledge and a large domain of skills that have been judged important for success in high school, college, and beyond. Thus, the Standards should be interpreted in a responsible way that will help students, parents, teachers, and administrators do the following:

- identify skill areas in which students might benefit from further instruction
- monitor student progress and modify instruction to accommodate learners' needs

- encourage discussion among principals, curriculum coordinators, and teachers as they evaluate their academic programs
- enhance discussions between educators and parents to ensure that students' course selections are appropriate and consistent with their plans for after high school
- enhance communication between secondary and postsecondary institutions
- identify the knowledge and skills that students entering their first year of postsecondary education should know and be able to do in English language arts, mathematics, and science
- assist students as they identify skills they need to master to prepare for college-level coursework

## 5.6 ACT College Readiness Benchmarks

### *5.6.1 Description of the ACT College Readiness Benchmarks*

The ACT College Readiness Benchmarks are the ACT scores that represent the level of achievement required for students to have a 50% chance of obtaining a B or higher or about a 75% chance of obtaining a C or higher in corresponding credit-bearing first-year college courses at a typical 2-year or 4-year postsecondary institution (Table 5.10). For example, the ACT English Benchmark (18) is the score associated with having a 50% chance of obtaining a B or higher in English Composition I.

**Table 5.10.** ACT College Readiness Benchmarks

| College course or course area | ACT test score | ACT Benchmark |
|---|---|---|
| English Composition I | English | 18 |
| College algebra | Math | 22 |
| American history, other history, psychology, sociology, political science, and economics | Reading | 22 |
| Biology | Science | 23 |
| Calculus I, biology, chemistry, physics, and engineering | STEM | 26 |
| English Composition I and social science courses | ELA | 20 |

Three separate studies were conducted to develop the current Benchmarks, and the reports documenting those studies provide more details on the study methodology and samples. The first study developed the ACT Benchmarks in English, mathematics, reading, and science (Allen, 2013). The second developed the STEM Benchmark (Radunzel et al., 2015), and the third developed the ELA Benchmark (Radunzel et al., 2017).

Benchmarks were developed for the courses or course combinations listed in Table 5.10. Success in a course was defined as earning a grade of B or higher in the course. Hierarchical logistic regression was used to model the probability of success in a course as a function of

ACT test score within each college. The student-level data were weighted to make the sample more representative of all ACT-tested students. For each course within each college, a cutoff score was chosen such that the probability of success (i.e., the probability of earning a B or higher in the course) was at least .50. This score point most accurately divided the sample into those who would be successful and those who would not (Sawyer, 1989b). The individual cutoff scores per college were weighted to make the sample more representative of all colleges with respect to institution type and selectivity (2-year, 4-year less selective, and 4-year more selective). The Benchmarks (Table 5.10) were determined by the median cutoff scores across colleges.

### *5.6.2 Intended Uses of the ACT College Readiness Benchmarks*

ACT recommends that the Benchmarks be used for any of three general purposes:

1. **Identifying students who are ready for credit-bearing courses or who need additional academic support.** Because success in college courses depends on more than just the knowledge and skills measured by the ACT test, the best course placement and early identification systems use multiple measures, such as high school GPA, ACT test scores, high school courses taken, and measures of social and emotional learning. The Benchmarks can be used to identify students who have the requisite knowledge and skills targeted by the ACT test. Because performance expectations and grading standards vary across colleges, the Benchmarks represent a standard for the typical postsecondary institution.

2. **Serving as a performance standard for K–12 students**. The Benchmarks can help states, districts, and schools identify the levels of performance on academic achievement tests that a student needs in order to be ready for college and career. The Benchmarks help articulate college expectations not only to students in high school but also to students in lower grades. Assessments designed for lower grades (e.g., PreACT®, PreACT® Secure™, and PreACT® 9 Secure™) can use the ACT test as the anchor of the assessment system and use the Benchmarks as the end target. Some states use the ACT test and the Benchmarks for federal or state accountability reporting.

3. **Monitoring educational improvement and achievement gaps over time.** Educational stakeholders at all levels (school, district, state, nation) are interested in how their institutions are improving and in how gaps between student groups change over time. The percentage of students meeting the Benchmarks can be used as one of the metrics for monitoring progress and setting goals, and it is most relevant when the ACT test is administered to all students. One advantage of using the Benchmarks for this purpose is that they are indicators of readiness for college coursework and so have relevance to students, educators, and policymakers.

### *5.6.3 Interpreting ACT Test Scores With Respect to Both the ACT College and Career Readiness Standards and the ACT College Readiness Benchmarks*

The performance levels on the ACT test necessary for students to succeed in college-level work are defined by the ACT College Readiness Benchmarks. Meanwhile, the knowledge and skills a student currently has (and areas for improvement) can be identified by examining the student's ACT test scores with respect to the ACT College and Career Readiness Standards. These two empirically derived metrics help students translate test scores into clear indicators of their current level of college readiness and identify key knowledge and skill areas that are needed to improve the likelihood of achieving college success.

# Chapter 6
# Scaling, Equating, and Technical Characteristics

This chapter describes the construction of score scales for the ACT test, the linking of the enhanced ACT, the procedures used to equate ACT test forms, and the psychometric properties of ACT scores. Scaling, linking, and equating the multiple-choice test sections are presented first, followed by scaling and equating the ACT writing test scores used to calculate the ELA score. The chapter also describes key psychometric characteristics of ACT scores, including reliability, measurement precision, and comparability across administrations and testing modes.

## 6.1 Scaling, Linking, and Equating the ACT English, Mathematics, Reading, and Science Test Sections

The ACT has maintained stable score scales since their establishment in 1989 to ensure comparability across administrations. Despite incremental changes to test specifications, all forms have been equated to these common scales. Before the formal launch of the enhanced ACT in 2025, a special linking study was conducted to align scores from the enhanced ACT with the legacy ACT score scales. Subsequent forms of the enhanced ACT are equated to the same scales, thereby preserving continuity in score interpretation over time.

### *6.1.1 Scaling*

The data used in the scaling process were collected in the fall of 1988 as part of the Academic Skills Study, which provided the basis for revising the ACT score scales and developing nationally representative norms. More than 100,000 high school students participated in the study. For scaling, a nationally representative sample of Grade 12 college-bound examinees was used. Sawyer (1989) discusses in detail the data used for scaling the ACT.

The scaling process consisted of three steps. First, weighted raw score distributions were computed for college-bound examinees from the Academic Skills Study. Second, the weighted raw score distributions were smoothed with a four-parameter beta compound binomial model (Kolen, 1991; Kolen & Hanson, 1989; Lord, 1965), and a double arcsine transformation was applied to equalize error variance across the score scale (Kolen, 1988). Finally, the smoothed and double arcsine transformed raw score distributions for Grade 12 college-bound examinees were linearly transformed to establish the ACT score scales. The second and third steps are described in greater detail in the rest of this section and by Kolen and Hanson (1989).

In the second step, smoothing the raw score distributions produced distributions that were easier to work with and that better estimated population distributions. Kolen (1991) and Hanson (1990) showed that smoothing techniques have the potential to improve the estimation of population distributions. Overall, the smoothing process resulted in distributions that appeared smooth without departing much from the unsmoothed distributions. In addition, the first three central moments (mean, variance, and skewness) of the smoothed distributions were identical to those of the original distributions. Values of the fourth central moment of the smoothed distributions (kurtosis) were either identical or very close to those of the original distributions.

The double arcsine transformation was applied to the smoothed raw scores to stabilize error variance. This ensured that the conditional standard error of measurement (*CSEM*) was approximately equal throughout the score scale for the Grade 12 college-bound examinees in the Academic Skills Study.

In the third step, initial scale scores were produced with a specified mean and a specified standard error of measurement (*SEM*). The process used methods introduced by Kolen (1988) and described in detail by Kolen and Hanson (1989). After a linear transformation was applied to set the mean score to 18 and the *SEM* to as close to 2 as possible, the initial scale scores were rounded to integers ranging from 1 to 36. Some adjustment of the rounded scale scores was then made to better meet the specified mean and *SEM* and to avoid gaps in the score scale (i.e., unused scale scores) or too many raw scores converting to a single scale score.

The score scales developed through this process range from 1 to 36 and are designed to have a constant *CSEM* of approximately 2. Since their establishment, scores from all test forms have been reported on these same scales through equating procedures. In cases where substantial changes to test specifications occur, statistical linking is used to maintain continuity with the established scales.

### *6.1.2 Linking*

As part of the transition to the enhanced ACT, a linking study was conducted in June 2024 to ensure the continuity of score interpretations between the legacy and enhanced test forms. The enhanced ACT introduces several changes, including reduced test length, increased time per item, embedded field-test items, an optional science section, and a revised Composite score definition based on only English, mathematics, and reading. Despite these modifications, the enhanced test measures the same constructs and reports scores on the existing 1–36 ACT scale. Li et al. (2025) provide a comprehensive description of the linking study. The study design and major findings are summarized in the rest of this section.

The linking study employed a random groups design in which examinees were randomly assigned to take either an enhanced or a legacy test form administered under comparable conditions. The analytic sample included approximately 6,900 students. Prior to linking, ACT evaluated construct comparability and test-taking behavior across forms. Results from confirmatory factor analyses indicated that the dimensional structure and factor loadings of the enhanced forms were similar to those of the legacy form, providing evidence of construct equivalence. In addition, analyses of completion rates, omit rates, rapid-response rates, and survey responses provided evidence of small positive effects on test completion and examinees' perceptions of having sufficient time to complete the test.

Item response theory (IRT) methods were used to link scores from the enhanced forms to the legacy ACT score scales. Item parameters were estimated for all study forms and transformed to the ACT scale using the Stocking–Lord procedure (Stocking & Lord, 1983). IRT true-score equating was used to link the scores of the enhanced forms to those of the legacy form. The results indicated that the psychometric properties of the enhanced forms were comparable to those of the legacy test and that the linking relationships were stable across subgroups.

Although the reduction in test length resulted in a small decrease in measurement precision, the effect on score interpretation was minimal.

Overall, the linking study findings support the conclusion that scores from the enhanced ACT can be reported on the legacy ACT scale and be interpreted in a manner consistent with scores from the legacy test. The study provided the basis for maintaining score comparability across test versions and for treating scores from the two versions as interchangeable for most intended uses. Subsequent forms of the enhanced ACT are equated and reported on the established score scales using the methodology described in the next section.

### *6.1.3 Equating*

New ACT test forms are developed each year. Although each form is constructed to adhere to the same content and statistical specifications, the forms may differ slightly in difficulty. To control for these differences, new forms are equated to a previously equated form. This process ensures that scores are reported on the same 1–36 ACT score scale and retain a consistent interpretation across all forms and test administrations.

ACT uses a random groups design to conduct equating. Examinees in the equating sample are assigned to spiraled test forms that include newly developed forms and a previously equated reference (anchor) form. The spiraling process ensures that randomly equivalent groups, typically exceeding 2,000 examinees per form, are administered each form.

IRT methodology is used to equate the enhanced forms. IRT equating begins with item calibration in BILOG-MG 3.0 (Zimowski et al., 2003) using data from ACT equating events. This involves estimating three parameters for each item: the difficulty parameter ($b_j$), the discrimination parameter ($a_j$), and the pseudo-chance (or "guessing") parameter ($c_j$), which represents the lower asymptote of the item characteristic curve and the probability that an examinee with very low ability will answer the item correctly. These three item parameters define the relationship between examinee ability, $\theta_i$, and the probability of responding correctly to an item in the 3-parameter logistic (3PL) IRT model (Birnbaum, 1968). The 3PL model is defined as

$$P\left(x_{ij} = 1 \middle| \theta_i, a_j, b_j, c_j\right) = c_j + \frac{1-c_j}{1+e^{-Da_j\left(\theta_i - b_j\right)}}, \tag{1}$$

where $x_{ij}$ is the item score (0 or 1) of examinee *i* to item *j*, *D* is a scaling constant equal to 1.7, and $\theta_i$ is the examinee latent trait (achievement in the case of the ACT).

The IRT parameters are then transformed to the same scale as all the previously calibrated items using the Stocking–Lord method (Stocking & Lord, 1983). Then a new form is equated to a reference form that has a preexisting raw-to-scale score conversion table. The equating process involves generating the test characteristic curves (TCCs) for the new form and the reference form. A TCC, which is simply the sum of the item characteristic curves (ICCs), shows the relationship between examinee ability and expected raw score (number of correct items) on a test form. In IRT true-score equating, scores on different forms are considered equivalent if they correspond to the same level of the latent trait ($\theta$). The equated raw scores are then used with the base form's raw-to-scale score conversion table to develop the raw-to-scale score conversion table for the new form.

The preceding discussion focused on equating the four multiple-choice ACT test sections. Other reported scores that are combinations of multiple test scores are not equated directly. These scores—the Composite, STEM, and ELA scores—are each a rounded average of the scale scores from two or more test sections. Chapter 5 provides more information on these scores. The Composite, STEM, and ELA scores are also comparable across forms because the scores used to compute them have been equated.

### 6.1.4 Score Scale Characteristics

The 1–36 ACT score scales were originally established to have a constant conditional standard error of measurement (*CSEM*) of about 2 across the score scale for each test section. For the enhanced ACT, this property continues to hold approximately, although reducing the number of items slightly increased the *CSEM*s, particularly for reading, where the *CSEM* is approximately 3. Having a constant *CSEM* allows for straightforward reporting of measurement precision across the score scale and supports score interpretation, consistent with the *Standards for Educational and Psychological Testing* (American Educational Research Association [AERA] et al., 2014, p. 39).

The approximately constant *CSEM* across the score scale allows an examinee's reported score to be interpreted as an estimate of the examinee's true score, with a typical measurement error of about ±1 *SEM*, which is approximately 2 scale score points for the English, mathematics, and science sections and approximately 3 scale score points for the reading section. With normally distributed error, about 68% of scores from repeated testing would fall within 1 *SEM* of the true score. Similarly, about 68% of examinees in a group would have observed scores within 1 *SEM* of their true scores. This characteristic supports the construction of simple confidence intervals and helps discourage overinterpretation of small score differences.

Although the ACT score scales were designed to yield approximately constant *CSEM*s, the *CSEM*s of operational forms may fluctuate over time. Therefore, ACT monitors and reports the *CSEM*s of newly administered forms each year. Section 6.3 provides additional information on *CSEM*s and other psychometric characteristics of ACT scores.

### 6.1.5 Reporting Category Readiness Ranges

As described in Chapter 3, ACT items are classified into reporting categories that describe specific groups of skills associated with college and career readiness. Student performance within a reporting category is reported as a percentage-correct score, which may fall within an ACT Readiness Range. The ACT Readiness Range indicates the score range expected of students who met or exceeded the corresponding ACT® College Readiness Benchmark (see Chapter 5 for a detailed description of reporting category scores). The ACT Readiness Range can vary across forms because of differences in difficulty and the number of items. The procedures used to establish the ACT Readiness Ranges are described next.

To determine the lower bound of a readiness range, student data are used to create a predictive relationship between ACT scale scores and percentage-correct scores in a reporting category. Using that relationship, the lower bound is set as the percentage-correct score expected for a student whose ACT scale score is equal to the corresponding ACT College Readiness Benchmark (e.g., 18 on English, 22 on mathematics). For example, a separate Readiness Range is developed for each of the three English reporting categories. For the Production of Writing reporting category, linear regression is used to estimate a predictive relationship between 1–36 English scale scores and percentage-correct scores on the items associated with the Production of Writing reporting category. This relationship is then used to identify the percentage-correct score for the reporting category corresponding to the Benchmark score of 18 on the overall English test. Students with percentage-correct scores at or above the lower bound are considered within the Readiness Range. The upper bound of each Readiness Range corresponds to examinees' answering all questions in that reporting category correctly. The same process is repeated to determine Readiness Ranges for the other two English reporting categories and the reporting categories of the other multiple-choice test sections.

## 6.2 Scaling and Equating the ACT Writing Test Section to Calculate ELA Scores

In September 2015, ACT launched the current ACT writing test section and then began reporting ELA scores. The ELA score is the rounded average of the English, reading, and writing scale scores. To support ELA score reporting, a 1–36 score scale was established for the writing test. Since September 2016, when the 2–12 rounded average domain scores replaced the 1–36 scores for reporting the ACT writing test score, the 1–36 writing scale scores have been used solely for the calculation of ELA scores.

The 1–36 writing scale was constructed based on data from the writing field-test study conducted in fall 2014. After all the prompts administered in the special study were evaluated, one prompt was selected as the base form for establishing the 1–36 writing score scale. To obtain the base prompt's raw-to-scale score conversion, percentile ranks were calculated for all raw score points (i.e., the sum of the four domain scores). Then the corresponding *z* scores from a standard normal distribution were obtained for these percentile ranks. The *z* scores were then linearly transformed to cover the whole score range of 1–36. Finally, a seventh-degree polynomial regression of the unrounded scale scores on the raw scores was used to slightly smooth the conversion before rounding to integer scale scores to obtain the final raw-to-scale score conversion for the base form.

Equating is conducted to ensure that the 1–36 writing scale scores will have a consistent interpretation across forms. Equipercentile equating with post-smoothing is used to equate the writing test scores. In equipercentile equating, a score on Form X and a score on Form Y are considered equivalent if the scores are associated with the same percentile rank for the randomly equivalent groups of examinees that took those forms (Kolen & Brennan, 2014). The equipercentile equating results are smoothed using an analytic method described by Kolen (1984) to establish a smooth relationship between scores on two test forms. The equivalent scores are then rounded to integers. The conversion tables resulting from this process are used to transform raw scores on the new forms to the 1–36 writing scale. This process ensures year-to-year comparability of ELA scores. Because the ELA score is based on the English, reading, and writing test sections, it generally provides a more reliable measure than the writing score alone, which is based on a response to a single writing prompt.

## 6.3 Reliability, Measurement Error, and Classification Consistency

The potential for some degree of inconsistency or error is inherent to the measurement of any cognitive characteristic. An examinee administered one form of a test on one occasion and a parallel form on another occasion may earn somewhat different scores on the two tests. These differences might be due to the examinee or the testing conditions, such as differential motivation or differential levels of distractions during the two administrations. These differences may also result from attempting to estimate the examinee's level of skill in a broad domain from a relatively small sample of items. This section describes the statistics used to quantify the reliability, measurement error, and classification consistency of ACT test scores.

### *6.3.1 Reliability and Standard Error of Measurement for the ACT Test Forms*

Reliability coefficients quantify the level of consistency in test scores across repeated administrations. These coefficients range from 0 to 1, with values closer to 1 indicating greater score consistency and those closer to 0 indicating little or no score consistency. Reliability coefficients are usually estimated based on a single test administration by calculating the inter-item covariances. Such coefficients are referred to as estimates of internal consistency reliability. Coefficient alpha (Cronbach, 1951), one of the most widely used estimates of internal consistency reliability, was computed for the ACT tests. Coefficient alpha can be computed using the formula

$$\alpha = \frac{k}{k-1}\left(1 - \frac{\sum_{i=1}^{k} s_i^2}{s_x^2}\right), \qquad (2)$$

where *k* is the number of test items, $s_i^2$ is the sample variance of the $i^{th}$ item scores, and $s_x^2$ is the sample variance of the observed total raw scores.

Coefficient alpha provides reliability estimates for number-correct scores. For scale scores, ACT uses the following formula to obtain a different reliability estimate ($r_t$):

$$r_t = 1 - \frac{SEM_t^2}{s_t^2}, \qquad (3)$$

where $SEM_t$ is the estimated scale score standard error of measurement for test *t*, and $s_t^2$ is the sample variance of the observed scale score for test *t*.

The standard error of measurement (*SEM*) summarizes the amount of measurement error or inconsistency in scores on a test. A four-parameter beta compound binomial model as described in Kolen, Hanson, and Brennan (1992) was used to estimate scale score reliability coefficients and *SEM*s. One input to these calculations was an estimate of the relative error variance from a generalizability study with a $person \times (items\colon content)$ design. Relative error variance is relevant when test scores are used to rank-order examinees. Reliability estimates based on absolute error variance, which is relevant to classification decisions, would be slightly lower (approximately 0.01 to 0.03 lower) than the reported values.

If normally distributed measurement errors are assumed, an interval extending 1 *SEM* above and below the reported score provides an approximately 68% confidence interval for the examinee's true score. That is, across repeated applications of this interval-construction procedure, approximately 68% of the resulting intervals would be expected to contain the corresponding true scores.

### Reliability and *SEM* for the ACT Test Scores

Scale score reliability estimates and *SEM* for the four multiple-choice ACT test section (English, mathematics, reading, and science), Composite, STEM, and ELA scores are shown in Table 6.1. These values were calculated using operational test data from eight of the test forms administered between June 2025 and April 2026. Reliability estimates were high, with values of 0.87 or greater for English, mathematics, Composite, STEM, and ELA scores and values of 0.81 or greater for reading and science scores. The *SEM* is about 1 for the Composite score, about 2 for the English, mathematics, and science scores, and close to 3 for reading scores. Reliability and *SEM* values were fairly consistent across forms.

**Table 6.1.** Summary Statistics of Scale Score Reliability and *SEMs* for the ACT Test Scores

| Test | No. of items | Reliability | | | SEM | | |
|---|---|---|---|---|---|---|---|
| | | Mdn | Min | Max | Mdn | Min | Max |
| **English** | 40 | 0.89 | 0.87 | 0.90 | 2.21 | 2.14 | 2.31 |
| **Mathematics** | 41 | 0.89 | 0.87 | 0.89 | 1.98 | 1.89 | 2.06 |
| **Reading** | 27 | 0.82 | 0.81 | 0.85 | 2.81 | 2.63 | 2.94 |
| **Science** | 34 | 0.84 | 0.82 | 0.86 | 2.21 | 2.10 | 2.35 |
| **Composite** | 108 | 0.94 | 0.94 | 0.95 | 1.36 | 1.32 | 1.41 |
| **STEM** | 75 | 0.92 | 0.92 | 0.94 | 1.48 | 1.45 | 1.54 |
| **ELA** | 68 | 0.91 | 0.90 | 0.92 | 1.61 | 1.57 | 1.64 |

*Note*. Mdn = median; Min = minimum; Max = maximum.

### Reliability and *SEM* for the ACT Reporting Category Scores

Raw score reliability (coefficient alpha) and *SEM*s were also calculated for the ACT reporting categories. These statistics, summarized in Table 6.2, were estimated using operational test data from eight forms administered between June 2025 and April 2026. Reliability estimates varied across reporting categories and were generally lower for categories with relatively few items. This result is expected because reliability is influenced by test length. Reporting category scores are not intended for high-stakes decision-making. Rather, they are provided to help identify students' strengths and weaknesses and to support instructional planning.

**Table 6.2.** Summary Statistics of Raw Score Reliability and *SEM* for the ACT Reporting Categories

| Test | Reporting categories | | No. of items | Reliability Mdn | Reliability Min | Reliability Max | *SEM* Mdn | *SEM* Min | *SEM* Max |
|---|---|---|---|---|---|---|---|---|---|
| English | Production of Writing | | 15–17 | 0.76 | 0.71 | 0.81 | 1.64 | 1.50 | 1.74 |
| | Knowledge of Language | | 7–9 | 0.64 | 0.57 | 0.71 | 1.11 | 1.02 | 1.17 |
| | Conventions of Standard English | | 16–17 | 0.79 | 0.75 | 0.80 | 1.73 | 1.64 | 1.83 |
| Math | **Preparing for Higher Mathematics** | Total | 33 | 0.86 | 0.84 | 0.89 | 2.45 | 2.36 | 2.51 |
| | | Number & Quantity | 4–5 | 0.46 | 0.35 | 0.56 | 0.96 | 0.87 | 0.97 |
| | | Algebra | 7–8 | 0.61 | 0.54 | 0.65 | 1.17 | 1.04 | 1.23 |
| | | Functions | 7–8 | 0.59 | 0.50 | 0.68 | 1.15 | 1.11 | 1.24 |
| | | Geometry | 7–8 | 0.56 | 0.51 | 0.69 | 1.18 | 1.10 | 1.23 |
| | Statistics & Probability | | 5–6 | 0.51 | 0.44 | 0.60 | 0.99 | 0.87 | 1.06 |
| | Integrating Essential Skills | | 8 | 0.63 | 0.57 | 0.67 | 1.21 | 1.05 | 1.25 |
| | Modeling | | 11–22 | 0.73 | 0.66 | 0.82 | 1.54 | 1.41 | 1.96 |
| Reading | Key Ideas & Details | | 12–13 | 0.70 | 0.67 | 0.77 | 1.44 | 1.41 | 1.56 |
| | Craft & Structure | | 8–9 | 0.57 | 0.48 | 0.63 | 1.27 | 1.17 | 1.35 |
| | Integration of Knowledge & Ideas | | 5–6 | 0.55 | 0.49 | 0.61 | 1.01 | 0.91 | 1.09 |
| Science | Interpretation of Data | | 13–16 | 0.71 | 0.62 | 0.75 | 1.59 | 1.44 | 1.68 |
| | Scientific Investigation | | 7–10 | 0.60 | 0.49 | 0.70 | 1.21 | 1.15 | 1.35 |
| | Evaluating Scientific Arguments and Models With Evidence | | 10–13 | 0.67 | 0.53 | 0.74 | 1.43 | 1.40 | 1.53 |

*Note*. Mdn = median; Min = minimum; Max = maximum; No. of items is the range based on the eight forms used in these analyses.

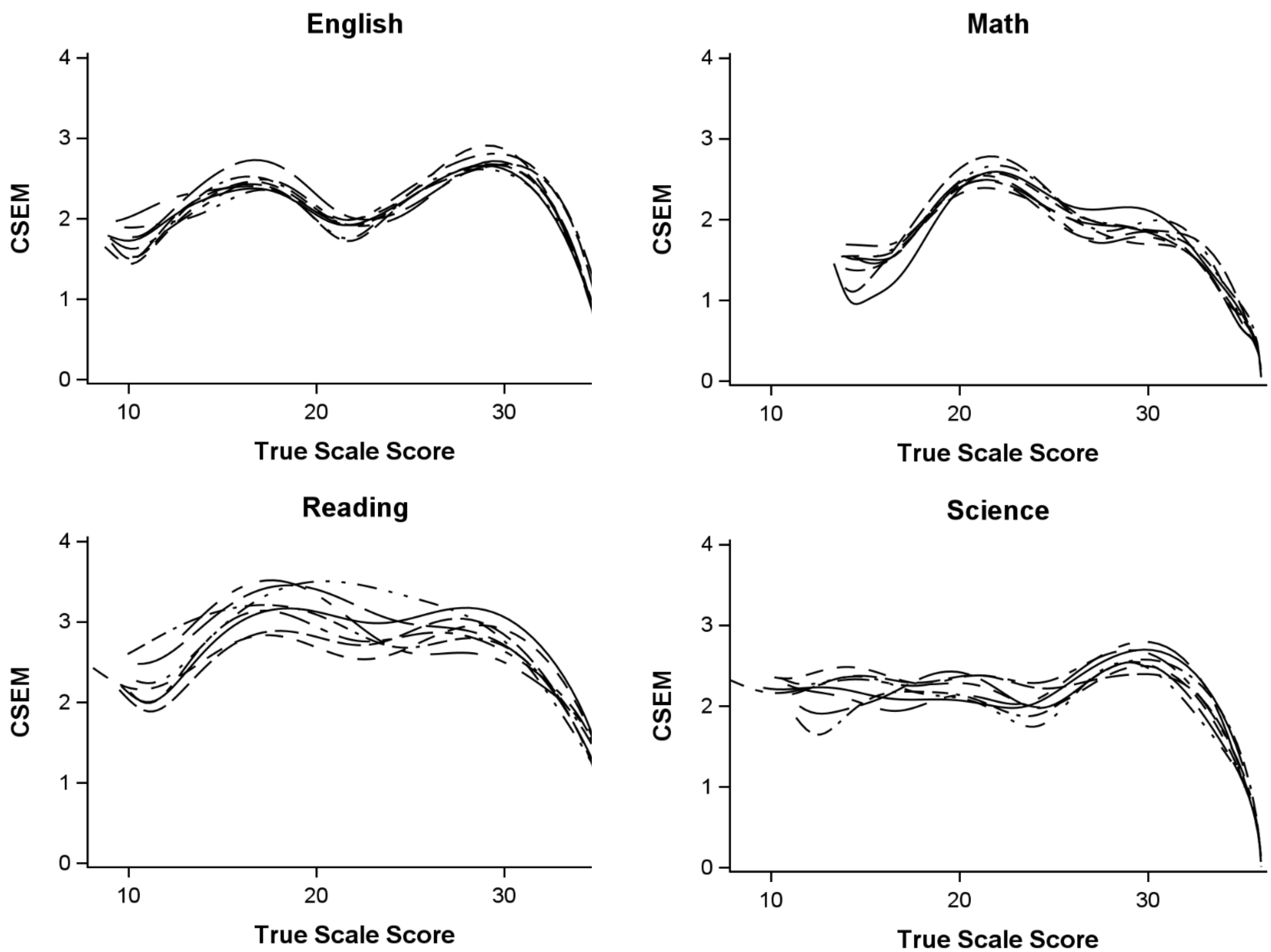
### Conditional Standard Errors of Measurement for ACT Multiple-Choice Test Scores

Whereas the *SEM* indicates average score uncertainty (or imprecision) across the entire score scale, the conditional standard error of measurement (*CSEM*) quantifies the uncertainty at a particular score. The ACT score scales were designed to have approximately constant *CSEM*s across the score scale. This statement implies, for example, that the *CSEM* for an ACT scale score is approximately the same for low-scoring examinees and high-scoring examinees.

For the ACT, *CSEM*s were computed using methods described by Kolen, Hanson, and Brennan (1992). Figure 6.1 presents the *CSEM*s for the four multiple-choice test sections for eight of the forms administered from June 2025 to April 2026. The *CSEM* is not plotted for very low scale scores that can be obtained by guessing or random responding. The minimum scale scores at which the *CSEM* was plotted were selected such that only an extremely small proportion of examinees would be expected to have a true scale score lower than the minimum plotted score.

The ACT test sections were scaled to have an approximately equal *CSEMs* along the score scales. This characteristic is most evident for the science test in Figure 6.1. The *CSEM*s of the English, mathematics, and reading test sections exhibit somewhat greater variation across the score scale than those for the science section. However, across most of the true scale score range, the *CSEM*s are generally close to approximately 2 scale score points for English, mathematics, and science and approximately 3 for reading. For all test sections, the *CSEM* approaches 0 as the true scale score approaches the maximum score of 36. For this reason, the *CSEM* cannot be perfectly constant for all true scale scores.

**Figure 6.1.** *CSEM* for Multiple-Choice Test Scores



### Reliability, *CSEM*, and Agreement Indices for ACT Writing Test Section Scores

ACT used the results from a generalizability study to estimate reliability and *CSEM*s for the ACT writing test section scores. To investigate the properties of the overall writing score and the domain scores, a generalizability study was conducted in fall 2014. The study was separated into three parts, each involving a different pair of schools. Within each pair of schools, two writing prompts were administered. The responses to both prompts were rated on the four writing domains by three raters. The same raters scored both prompts within each school pair. Each pair of schools had different pairs of prompts and different groups of three raters, which essentially served as three replications of the same study.

Across all three school pairs, the estimated variance components for the rater-by-prompt and rater-by-person (or student) interactions were small, indicating that raters scored prompts similarly and that students received generally consistent ratings from the different raters. In contrast, the estimated variance component for the person-by-prompt interaction was relatively large for all three pairs of schools. This finding was consistent with results typically observed in the research literature on extended-response assessments. For the average of the writing domain scores, the generalizability coefficients (reliability-like estimates of score consistency) ranged from 0.61 to 0.77, which are fairly high for a writing assessment. Corresponding *SEM*s ranged from 0.84 to 1.10.

Data from the 2019 writing field test study were used to estimate the reliability and *SEM* for writing scores on the 1–36 scale used in the calculation of ELA scores. Each student responded to two different prompts, and the data were analyzed using a person-by-occasion generalizability study design. The individual conditional error variances were fit with a fifth-degree polynomial. The square roots of these fitted conditional error variances are represented by the solid line in Figure 6.2. The average *CSEM* values, represented by the circles, were calculated by taking the square root of the average conditional error variances at each scale score point. The generalizability coefficient was 0.74, and the scale score *SEM* was 3.23. This *SEM* estimate was used to calculate the reliability and *SEM* of the ELA score.

**Figure 6.2.** Average and Fitted *CSEM*s for ACT Writing Test Section Scale Scores

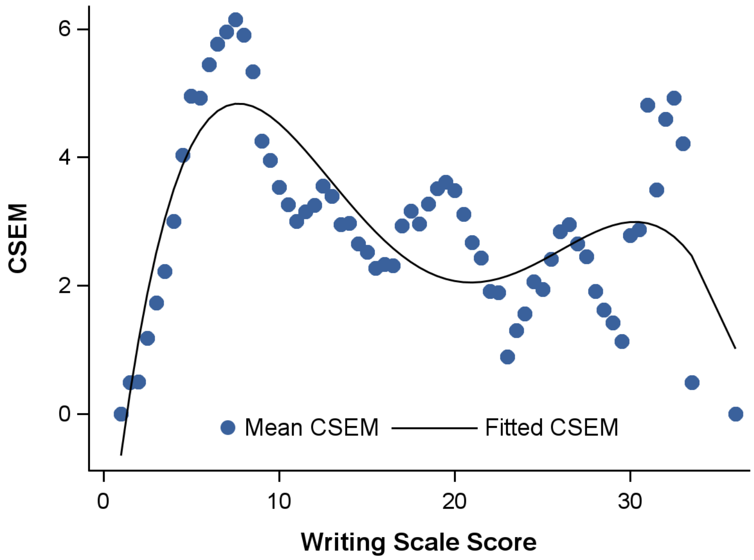



Table 6.3 summarizes operational rater agreement indices for the ACT writing domain scores. Median, minimum, and maximum values are reported across eight forms administered between June 2025 and April 2026 for each domain and agreement index. The agreement indices include the perfect agreement rate, the perfect-plus-adjacent agreement rate, and the quadratic weighted kappa coefficient.

The perfect agreement rate, defined as the proportion of examinees who received the same domain score (from 1 to 6) from two raters, ranged from 0.657 to 0.816 across domains and forms. The perfect-plus-adjacent agreement rates, defined as the proportion of examinees who received either the same domain score or adjacent domain scores (e.g., 5 and 6) from both raters, were very high, ranging from 0.989 to 0.999 across domains and forms.

The quadratic weighted kappa coefficient (Cohen, 1968) is a measure of agreement between the raters for categorical scores (e.g., 1, 2, 3). It uses weights to account for the relative differences between categories. In the calculation, for example, a 2-point disagreement is weighted more than a 1-point disagreement. The kappa coefficient is positive when the observed agreement is larger than the level of agreement attributable to chance, with larger numbers representing stronger agreement between two raters. Fleiss, Levin, and Paik (2003) suggested that for most purposes, kappa values larger than 0.75 represent excellent agreement beyond chance, values below 0.40 represent poor agreement beyond chance, and values in between represent fair to good agreement beyond chance. The quadratic weighted kappa coefficients for the ACT writing domain scores ranged from 0.716 to 0.876, indicating good rater agreement.

**Table 6.3.** Agreement Rates for the ACT Writing Domain Scores

| Domain | Agreement index | Median | Minimum | Maximum |
|---|---|---|---|---|
| **Ideas & Analysis** | Perfect agreement | 0.737 | 0.665 | 0.788 |
| | Perfect + adjacent agreement | 0.996 | 0.992 | 0.998 |
| | Quadratic weighted kappa | 0.811 | 0.771 | 0.859 |
| **Development & Support** | Perfect agreement | 0.738 | 0.659 | 0.816 |
| | Perfect + adjacent agreement | 0.997 | 0.991 | 0.999 |
| | Quadratic weighted kappa | 0.813 | 0.775 | 0.876 |
| **Organization** | Perfect agreement | 0.740 | 0.666 | 0.795 |
| | Perfect + adjacent agreement | 0.997 | 0.992 | 0.999 |
| | Quadratic weighted kappa | 0.809 | 0.758 | 0.855 |
| **Language Use & Conventions** | Perfect agreement | 0.746 | 0.657 | 0.776 |
| | Perfect + adjacent agreement | 0.997 | 0.989 | 0.999 |
| | Quadratic weighted kappa | 0.746 | 0.716 | 0.829 |

### *CSEM* for Composite Scores

Assuming that measurement errors on the ACT multiple-choice test sections (English, mathematics, and reading) are independent, the *CSEM* for the unrounded Composite score is given by Equation 4,

$$s_c(\tau_e, \tau_m, \tau_r) = \frac{\sqrt{\sum_i s_i^2(\tau_i)}}{3}, \tag{4}$$

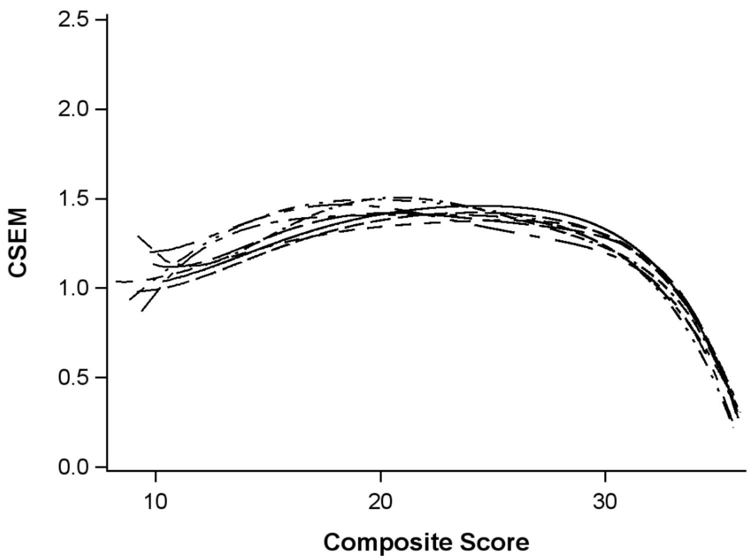
where $s_i(\tau_i)$ is the *CSEM* for test *i* at true scale score $\tau_i$ and *i* = *e*, *m*, and *r* denote the English, mathematics, and reading test sections, respectively. The $s_i(\tau_i)$ functions are plotted in Figure 6.1.

To generate *CSEM* plots for Composite scores, observed proportion-correct scores (i.e., the number of items answered correctly divided by the total number of items) for examinees on the three test sections were treated as true proportion-correct scores at which the *CSEM*s were calculated. For each section, the *CSEM* was computed for each examinee using the observed proportion-correct score as the true proportion-correct score in the formula for the *CSEM* (Equation 8 in Kolen et al., 1992). In addition, for each test section, the true scale score corresponding to the observed proportion-correct score (treated as a true proportion-correct score) was computed (Equation 7 in Kolen et al., 1992). The resulting *CSEM*s for the three test sections were substituted in Equation 4 to compute the Composite score *CSEM*.

A particular true Composite score can be obtained from multiple combinations of true scores on the three test sections. Consequently, each true Composite score may correspond to multiple *CSEM* values, depending on the combination of English, mathematics, and reading true scores that produced that Composite score. To obtain a unique *CSEM* value at each Composite score for each test form, a fifth-degree polynomial regression was fitted to the Composite score *CSEM*s. Figure 6.3 plots the fitted *CSEM*s for the Composite score for eight test forms administered between June 2025 and April 2026. The Composite score *CSEM*s were reasonably constant across the score scale.

A limitation of the approach used to produce the Composite score *CSEM*s shown in Figure 6.3 is that the estimates correspond to the unrounded average of the three test section scores rather than the rounded average of the three test section scores, which is the Composite score reported to examinees.

**Figure 6.3.** *CSEM* for Composite Scores




### *CSEM* for STEM and ELA Scores

The *CSEM*s for the STEM and ELA scores were calculated using the same approach used to calculate the Composite score *CSEM*s. Assuming that measurement errors on the four multiple-choice test sections are independent, the *CSEM* for the unrounded STEM score is

$$S_{STEM}(\tau_m, \tau_s) = \frac{\sqrt{\sum_i s_i^2(\tau_i)}}{2}, \quad (5)$$

where *i* = *m* and *s* for mathematics and science, respectively. Similarly, the *CSEM* for the unrounded ELA scores is

$$S_{ELA}(\tau_e, \tau_r, \tau_w) = \frac{\sqrt{\sum_i s_i^2(\tau_i)}}{3}, \quad (6)$$

where $s_i(\tau_i)$ is the *CSEM* for test *i* at true scale score $\tau_i$ and *i* = *e*, *r*, and *w* for English, reading, and writing, respectively. The same data used to produce the Composite Score *CSEM* values were used to obtain the *CSEM* values for the STEM scores plotted in Figure 6.4 and the *CSEM* values for the ELA scores in Figure 6.5.

**Figure 6.4.** CSEM for STEM Scores

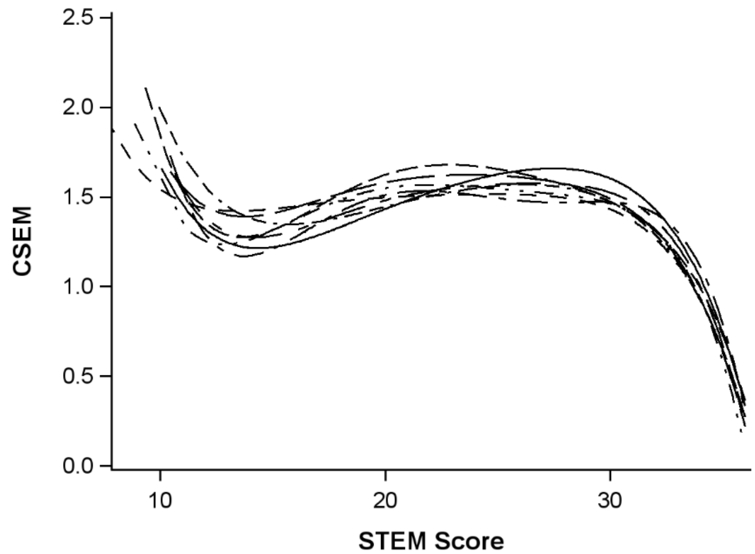
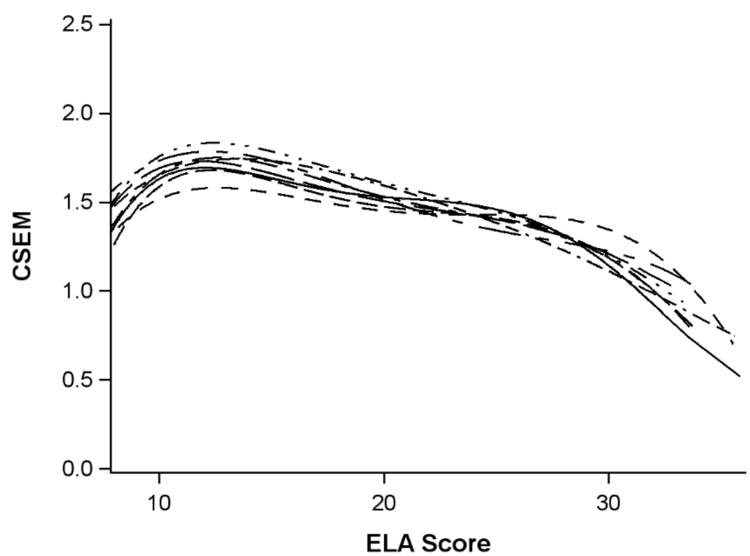



**Figure 6.5.** *CSEM* for ELA Scores




### *6.3.2 Classification Consistency*

Classification consistency refers to the extent to which examinees are classified into the same category over replications of a measurement procedure. Because tests are rarely administered twice to the same examinee, classification consistency is typically estimated from a single test administration using statistical models that make strong assumptions about the distributions of true scores and measurement errors (e.g., Hanson & Brennan, 1990; Livingston & Lewis, 1995).

Using the method described by Livingston and Lewis (1995), the true score distribution was estimated by fitting a four-parameter beta distribution. The expected conditional distribution of observed scores, given the true score, is a binomial distribution. Assuming independent errors of measurement, the probabilities of classification into each possible pair of categories were computed for each true score. These conditional probabilities were then aggregated over the true score distribution to produce a contingency table containing the probabilities of a student receiving scores from two administrations that fall into any combination of categories. The estimated classification consistency index for the whole group is the sum of the values on the diagonal of the contingency table, which represent the probabilities of being classified in the same category on two separate administrations. The next section gives the classification consistency results for the ACT test scores and indicators.

### Classification Consistency for the ACT Multiple-Choice Test, STEM, and ELA Scores

Data from eight forms administered from June 2025 to April 2026 were used to compute classification consistency values for the four ACT multiple-choice test sections and the STEM and ELA scores. Classifications were based on the ACT College Readiness Benchmarks (see Chapter 5 for details). Table 6.4 summarizes the classification consistency results. The values are all fairly high, ranging from a low of 0.82 for science to a high of 0.92 for STEM.

**Table 6.4.** Classification Consistency for the ACT College Readiness Benchmarks

| Test | Number of Items | Median | Minimum | Maximum |
|---|---|---|---|---|
| **English** | 40 | 0.87 | 0.86 | 0.89 |
| **Mathematics** | 41 | 0.87 | 0.86 | 0.88 |
| **Reading** | 27 | 0.84 | 0.83 | 0.86 |
| **Science** | 34 | 0.84 | 0.82 | 0.85 |
| **STEM** | 75 | 0.91 | 0.89 | 0.92 |
| **ELA** | 68 | 0.87 | 0.86 | 0.88 |

Similarly, classification consistency was computed for the ACT Readiness Ranges for each ACT reporting category. Table 6.5 provides these estimates that are based on data from eight of the forms administered from June 2025 to April 2026. As expected, reporting categories with fewer items generally exhibited lower classification consistency.

**Table 6.5.** Classification Consistency for the ACT Readiness Ranges

| Test | Reporting categories | | No. of items | Median | Min | Max |
|---|---|---|---|---|---|---|
| **English** | Production of Writing | | 15–17 | 0.83 | 0.81 | 0.86 |
| | Knowledge of Language | | 7–9 | 0.80 | 0.74 | 0.85 |
| | Conventions of Standard English | | 16–17 | 0.81 | 0.77 | 0.88 |
| **Math** | **Preparing for Higher Mathematics** | — | 33 | 0.84 | 0.83 | 0.87 |
| | | Number & Quantity | 4–5 | 0.65 | 0.61 | 0.72 |
| | | Algebra | 7–8 | 0.73 | 0.70 | 0.75 |
| | | Functions | 7–8 | 0.72 | 0.67 | 0.75 |
| | | Geometry | 7–8 | 0.69 | 0.66 | 0.78 |
| | | Statistics & Probability | 5–6 | 0.70 | 0.63 | 0.78 |
| | Integrating Essential Skills | | 8 | 0.73 | 0.70 | 0.75 |
| | Modeling | | 11–22 | 0.78 | 0.74 | 0.82 |
| **Reading** | Key Ideas & Details | | 12–13 | 0.77 | 0.75 | 0.83 |
| | Craft & Structure | | 8–9 | 0.73 | 0.70 | 0.75 |
| | Integration of Knowledge & Ideas | | 5–6 | 0.70 | 0.63 | 0.76 |
| **Science** | Interpretation of Data | | 13–16 | 0.77 | 0.75 | 0.79 |
| | Scientific Investigation | | 7–10 | 0.72 | 0.65 | 0.78 |
| | Evaluating Scientific Arguments and Models With Evidence | | 10–13 | 0.76 | 0.68 | 0.80 |

### Classification Consistency for the Progress Toward WorkKeys NCRC Indicator

The Progress Toward the ACT® WorkKeys® National Career Readiness Certificate™ (NCRC®) had classification consistency values ranging from 0.72 to 0.79 across eight of the forms administered from June 2025 to April 2026. These values are quite high considering that there are five performance levels for the WorkKeys NCRC, as shown in Table 6.6. Note that the classification consistency index is an indication of the stability of the Progress Toward the WorkKeys NCRC Indicator if different ACT test forms were taken and is not an indication of the accuracy of the classification compared with students' actual NCRC attainment. See Chapter 5 for more information about the Progress Toward the ACT WorkKeys National Career Readiness Certificate Indicator.

**Table 6.6.** Composite Score Ranges for the WorkKeys NCRC Levels

| Most likely WorkKeys NCRC level | Composite score range |
|---|---|
| None | 1–12 |
| Bronze | 13–15 |
| Silver | 16–20 |
| Gold | 21–26 |
| Platinum | 27–36 |

## 6.4 Mode Comparability

### *6.4.1 Overview of ACT Test Administration Modes*

The ACT may be administered on paper or online for National and for State and District testing, and it is administered only online for International testing.

In April 2014, ACT conducted a pilot study of online ACT testing, administering the ACT test to approximately 4,000 students at 80 test sites and reporting college-reportable scores. Online ACT testing expanded to a limited number of U.S. test sites for State and District administrations in April 2015. Beginning in 2016, online ACT testing became available for State and District administrations, and in September 2018, all International ACT administrations transitioned to online delivery. Online ACT testing was subsequently introduced for the National testing program in February 2024.

Following a pilot online study for the enhanced ACT in June 2024, ACT introduced enhanced forms of the ACT test for National online testing in April 2025. A full rollout to both paper and online administrations for National testing occurred in September 2025, followed by implementation for State and District and International online administrations in February 2026.

### *6.4.2 Online Platform and Capabilities*

For National testing, ACT uses the PSI Universal Test Driver (UTD) platform. Test centers offering online testing may provide test center–managed devices, and testing is currently supported on Chromebooks, Windows desktops and laptops, and macOS devices. Beginning with the September 2025 National test administration, ACT introduced the option to bring your own device (BYOD) that allows students to test on their own computers, if approved. Additional information about BYOD is available at ACT Gateway: Student Resources on the ACT website.

For State and District testing, ACT uses the TestNav platform for online test delivery. The TestNav platform supports testing on multiple device types, including desktop and laptop computers running macOS, Microsoft Windows, and Chrome OS. ACT regularly publishes the technical requirements of devices to maintain compatibility with evolving testing technologies. Current technical requirements and related guidance for ACT State and District online testing are available in the Technical Guide for Online Testing.

International online ACT administrations are delivered through the PSI UTD platform. International test centers can administer the ACT on desktop and laptop computers running

Microsoft Windows or macOS. Current requirements are available at Technical Requirements: Information for International CBT Testing Technical Coordinators.

### 6.4.3 Comparability of Scores Between Online and Paper Testing

When a test is offered across multiple modes, its scores must support the same intended interpretations and uses. The *Standards for Educational and Psychological Testing* (AERA et al., 2014) require that programs offering multiple delivery modes provide a sound rationale and empirical evidence that score validity is not compromised by the mode of delivery.

When mode effects are detected—that is, when examinees of equivalent ability perform systematically differently across modes—equating adjustments are necessary to ensure that scores remain interchangeable. ACT maintains score comparability between online and paper administrations by conducting mode comparability studies and subsequent form equating. ACT has conducted a series of such studies at each major transition in delivery platform and test design to evaluate and address potential mode effects.

#### Mode Comparability Studies: Overview

As part of the initial development of online ACT testing, ACT conducted several special studies to evaluate the comparability of scores between online and paper administrations prior to the official launch of the ACT online tests in the State and District, National, and International administrations, including a timing study in fall 2013 and mode comparability studies in spring 2014 and spring 2015. In 2018, another mode comparability study was conducted in preparation for online testing for the ACT International program. Then, between 2019 and 2020, three mode comparability studies were conducted to evaluate and support the use of the TAO platform for online ACT delivery. A mode study was then conducted in December 2023 to support the launch of online testing for the National testing program using the PSI UTD platform. Finally, following the introduction of the enhanced ACT and the June 2024 linking study, ACT conducted a mode comparability study in October 2024 to evaluate score comparability between the paper and online administrations of the enhanced ACT (ACT, 2026). All studies except the December 2023 mode study used a randomly equivalent groups design; that is, students were randomly assigned to take the test under different timing conditions in the online timing study, and they were randomly assigned to take the paper or online test in the mode studies.

#### 2013–2015 Pearson TestNav Mode Comparability Studies

To support the transition to online testing, a series of studies were conducted between 2013 and 2015 to evaluate timing requirements and score comparability between paper and online administrations of the ACT (Li et al., 2017). These studies implemented randomized experimental designs, which are widely considered the gold standard for evaluating timing and mode effects.

The fall 2013 timing study included approximately 3,000 students from 58 schools who were randomly assigned to different timing conditions. Results indicated that the standard paper time limits were appropriate for the English and mathematics sections, whereas the online reading and science sections initially appeared more speeded, as reflected in higher item omission rates and student survey responses. Based on these findings, 5 minutes were provisionally added to the online reading and science tests for subsequent evaluation.

The spring 2014 and 2015 mode comparability studies randomly assigned students to paper and online testing conditions in operational ACT administrations. Across both studies, online examinees scored slightly higher than paper examinees. The largest differences were observed for English and reading. In 2014, the effect sizes were 0.15 for English and 0.32 for reading; in 2015, they were 0.15 for English and 0.14 for reading. The mathematics section showed little to no meaningful mode effect.

Additional analyses demonstrated strong construct comparability between paper and online administrations. Comparisons of factor structures, inter-test correlations, effective weights, and item response theory (IRT) parameters indicated that both modes measured the same underlying academic constructs with similar levels of measurement precision. Although small differences were observed in some item- and test-level statistics, the largest mean differences under the final online timing conditions were approximately 1 ACT scale score point for English and reading (effect sizes of about 0.17–0.18 standard deviations).

Because ACT scores are used for high-stakes purposes, ACT applied equipercentile equating procedures to produce mode-specific score conversion tables and ensure score comparability across testing modes. After equating, score distributions were comparable across modes. The 2015 study also demonstrated that the additional time provided for online reading and science was unnecessary, leading ACT to reinstate the standard timing limits. The spring 2015 online form served as the base online form linked to paper testing, and subsequent online forms used in State and District testing programs have been equated to this form to maintain comparability regardless of testing mode.

### 2018 PSI ATLAS Cloud Mode Comparability Study

In preparation for expanding online testing to the International program in 2018, a mode comparability study was conducted in which participants were randomly assigned to paper or online testing on the ATLAS Cloud platform.

Analyses paralleled those conducted in the TestNav studies. Results indicated very small differences between paper and online testing in correlations among tests, effective weights, and reliability. However, average scores were slightly higher for online examinees, particularly in English and reading. The average differences (in standard deviation units) were approximately 0.16 (English), 0.05 (mathematics), 0.24 (reading), and 0.07 (science).

Overall, results suggested that mode effects on ATLAS Cloud were similar in magnitude to those observed on TestNav. On the writing test, scores also tended to be higher for online examinees. ACT applied equating procedures to adjust for these differences and ensure comparability of reported scores.

### 2019–2020 TAO Mode Comparability Studies

To prepare for expanding online National testing to the TAO platform, ACT conducted three mode comparability studies in October 2019, December 2019, and February 2020 (Steedle et al., 2020). Although ACT no longer uses TAO for National testing, the results from the three mode studies are relevant. Each study administered a different ACT form using a randomly equivalent groups design, and all participants received college-reportable scores.

Findings were consistent across the three studies and also with prior ACT mode studies: online examinees scored modestly higher than paper examinees on English ($d$ = 0.10–0.13) and reading ($d$ = 0.16–0.22); mathematics effects were negligible; and science effects were small to moderate ($d$ = 0.04–0.12). Item-level analyses suggested that these differences were primarily attributable to differential speededness, particularly in the latter portions of the English and reading test sections. The February 2020 study also found a writing mode effect ($d$ = 0.39). Confirmatory factor analyses supported construct equivalence across modes in all three studies.

In all cases, ACT equated online forms to paper forms to ensure that reported scores were comparable regardless of testing mode.

### December 2023 PSI UTD Mode Study (National Testing)

ACT conducted a mode comparability study in December 2023 to support the February 2024 launch of online National testing on the PSI UTD platform. Unlike the prior randomly equivalent groups studies, this study employed a quasi-experimental matched-samples design, using propensity score matching to compare the scores of paper and online examinees.

Results were consistent with findings from earlier mode studies. Differences between modes were minimal for correlations among test sections, effective weights, internal consistency reliability, and confirmatory factor analysis results. However, some test sections, particularly reading and writing, showed notable mean differences. Effect sizes of online test scores minus paper test scores were 0.01 for English (nonsignificant), −0.03 (nonsignificant) for mathematics, 0.23 for reading, 0.10 for science, and 0.37 for writing.

As with its prior studies, ACT applied equating procedures to ensure that scores from online and paper administrations were comparable.

### October 2024 Mode Comparability Study for the Enhanced ACT Test

Following the June 2024 linking study, a mode comparability study was conducted in October 2024 to evaluate whether scores from paper and online administrations of the enhanced ACT were equivalent and to support interchangeable score reporting across modes (ACT, 2026).

The October 2024 study was conducted in an operational testing environment in which participants received college-reportable scores. Data were collected during the October National and State and District (school-day) administration using a random groups design. Students in selected national test centers were randomly assigned to either paper testing or online testing administered through the UTD platform, and students in participating school-day settings were randomly assigned to either paper testing or online testing administered through the TestNav platform. The final cleaned dataset included 4,791 students (ACT, 2026).

Analyses were conducted to evaluate whether the delivery platform influenced examinee experience, item performance, or score comparability. Results showed no meaningful platform effects; therefore, data from the State and District and the National administrations were combined for subsequent analyses (ACT, 2026).

The study evaluated two primary aspects of mode comparability: score equivalency and construct equivalency. Score equivalency pertains to the similarity of observed score distributions across paper and online administrations for randomly equivalent groups, whereas construct equivalency pertains to whether the two modes measure the same underlying abilities. Analyses were conducted at both the test and item levels for the English, mathematics, reading, and science sections, including comparisons of score distributions, correlations among section tests, reliability estimates, confirmatory factor analysis (CFA) results, and differential item functioning (DIF) across modes (ACT, 2026).

At the test level, raw score and scale score means were compared across modes. Scale score comparisons were based on applying the online conversions established in the June 2024 linking study to both the online and paper groups (ACT, 2026; Li et al., 2025).

Prior to mode adjustment, mean scores were slightly higher for online testing than for paper testing across all test sections, with the largest differences observed for reading and the smallest for science. Mean scale score differences were 0.578 for English, 0.396 for mathematics, 0.922 for reading, and 0.217 for science, corresponding to small effect sizes of 0.096, 0.082, 0.140, and 0.041, respectively. Differences were statistically significant for English, mathematics, and reading but not for science. Distributional analyses yielded consistent results. Score distributions for mathematics and science were highly similar across modes, whereas English and reading scores tended to be higher for the online group. Tests of distributional equivalence indicated statistically significant differences for English and reading (ACT, 2026).

Evidence supporting construct equivalency was strong. Correlations among section scores, effective weights, and Cronbach's α reliability estimates were highly similar across modes. Reliability estimates were nearly identical for paper and online administrations, indicating comparable internal consistency across modes (ACT, 2026). Confirmatory factor analyses demonstrated acceptable model fit for all test sections across both modes, with similar factor loadings observed for paper and online administrations. DIF analyses identified only a small number of items with statistically detectable differences, and all such items were judged to function appropriately and measure the intended constructs across modes (ACT, 2026).

Item-level analyses indicated that items were generally easier for online examinees, particularly on the reading section and for items appearing later in the test. Item omission rates were consistently lower for online administrations in the later portions of the test sections, which partially accounted for the higher item difficulty estimates under online conditions. Differences in omission rates were smallest for mathematics (ACT, 2026).

To support comparable scores across modes, ACT used IRT true-score equating to apply a mode adjustment. Item parameters from paper testing were transformed to the online scale using the Stocking and Lord (1983) method, and IRT true-score equating was conducted to obtain score conversion tables for paper administrations. After mode adjustment, scale score distributions were nearly identical across paper and online administrations for all four test sections, supporting the conclusion that reported scores are comparable regardless of delivery mode (ACT, 2026).

Findings from the October 2024 study were consistent with prior ACT mode comparability research, which has generally shown small positive mode effects favoring online testing, particularly for English and reading, with minimal differences for mathematics and science (Arthur et al., 2020; Steedle et al., 2020; Wang & Steedle, 2020).

Taken together, the October 2024 findings indicate that paper and online administrations of the enhanced ACT measure the same constructs and yield highly similar score distributions. Small pre-adjustment differences were effectively addressed through equating procedures, ensuring that scores from both modes are interchangeable.

# Chapter 7
# Validity Evidence for the ACT Tests

According to the *Standards for Educational and Psychological Testing* (American Educational Research Association [AERA] et al., 2014), “validity refers to the degree to which evidence and theory support the interpretations of test scores for proposed uses of tests” (p. 11). Arguments for the validity of an intended inference made from a test score may contain logical, empirical, and theoretical components. A distinct validity argument is needed for each intended use of a test score.

Consistent with the *Standards*, validation is an ongoing process. Validity is established through the accumulation of evidence over time rather than by any single study, and the body of evidence supporting a given use is expected to grow and be refined as new data become available (AERA et al., 2014). Much of the evidence summarized in this chapter was gathered using the ACT® test prior to the enhancements introduced in spring 2025. ACT is actively accumulating additional validity evidence for the enhanced ACT. Allen et al. (2026) documents the design rationale, the content and format changes, and the plan for gathering validity evidence for the enhanced ACT. Readers seeking further detail should consult that document. Initial evidence based on the enhanced ACT data is summarized in Section 7.1.13 of this chapter, with additional evidence to be added as it becomes available.

The potential interpretations and uses of ACT test scores are numerous and diverse, and each needs to be justified by a validity argument. This chapter describes content, construct, and criterion validity evidence for five of the most common interpretations and uses: measuring students’ educational achievement in particular subject areas, making college admissions decisions, making college course placement decisions, evaluating students’ likelihood of success in the first year of college and beyond, and evaluating program effectiveness.

## 7.1 Using ACT Scores to Measure Educational Achievement

ACT tests measure students’ problem-solving skills and knowledge in particular subject areas. The usefulness of ACT scores for this purpose provides the foundation for validity arguments for more specific uses (e.g., course placement). This section comprises 13 subsections and provides validity evidence for using ACT test scores to measure students’ educational achievement. The first subsection summarizes content validity evidence supporting the interpretation of ACT scores as a measure of educational achievement. The second covers evidence from cognitive lab studies. The next five focus on relating high school coursework, grades, end-of-course exam scores, and noncognitive factors to ACT scores and ACT® College Readiness Benchmark attainment. The eighth subsection focuses on understanding subgroup differences on the ACT. The ninth focuses on the relationships between test preparation activities and ACT performance. The tenth addresses the use of ACT scores for measuring educational achievement for gifted and talented programs. The eleventh describes validity evidence related to the interpretation of scores for examinees who use English learner supports during the test. The twelfth examines grade inflation and its implications for ACT score interpretation and use. The final subsection summarizes validity evidence supporting the enhanced ACT.

### *7.1.1 Content-Oriented Evidence for ACT Scores*

The guiding principle underlying the development of the ACT is that the best way to predict success in college is to measure as directly as possible the degree to which each student has developed the academic skills and knowledge that are important for success in college. Tasks presented in the tests must therefore be representative of scholastic tasks. They must be intricate in structure, comprehensive in scope, and significant in their own right, rather than narrow or artificial tasks that can be defended for inclusion on the tests solely because of their statistical correlation with a criterion. In other words, assessment tasks must be designed to match the content and cognitive demands of the associated academic domain. Thus, content-related validity is particularly significant in this context.

The ACT tests contain a proportionately large number of complex problem-solving exercises and few measures of narrow skills. The tests are oriented toward major areas of high school and college instructional programs. Thus, ACT scores and the skill statements based on the ACT® College and Career Readiness Standards are directly related to student educational progress and can be readily understood and interpreted by instructional staff, parents, and students.

As described in Chapters 2 and 3, the test development procedures include an extensive review process, with each item being critically examined at least 16 times. Detailed test specifications have been developed to ensure that the test content represents current high school and college curricula. All test forms are reviewed to ensure that they match these specifications. Hence, there is an ongoing evaluation of the content validity of the tests during the development process.

The standardization of ACT tests is also important to their proper use as measures of educational achievement. Because ACT scores have the same meaning for all students, test forms, and test dates, they can be interpreted without reference to these characteristics.[2] The courses students take in high school and the grades they earn are also measures of educational achievement, but these variables are not standardized because course content varies considerably among schools, and grading policies vary among instructors. Therefore, while high school courses taken and grades earned are measures of educational achievement, their interpretation should properly take into account differences in high school curricula and grading policies. ACT scores, because they are standardized measures, are more easily interpreted for the purpose of comparing students than are courses taken and grades earned.

[2] ACT scores obtained before October 1989, however, are not directly comparable to scores obtained in October 1989 or later owing to the release of a new version of the ACT in October 1989. Although scores on the current and former versions are not directly comparable, approximate comparisons can be made using a concordance table developed for this purpose (ACT, 1989).

Additionally, external expert panels were conducted to validate recent enhancements to the ACT. Experts in ELA, math, and science were provided with the test construct, the test blueprints, sample items, sample forms, and domain sampling plans. Across all domains, the experts concurred that the proposed enhancements met the goals for evaluating college and career readiness and assessing whether students met common state standards in ELA, math, and science. Experts across all content areas stated that the items were able to elicit multistep critical thinking processes. Full details concerning the external expert panels can be found in Chapter 2 of the *Design Framework for the ACT Enhancements* (Allen et al., 2026).

### *7.1.2 Evidence From Cognitive Lab Studies*

Cognitive lab studies involve think-aloud protocols, wherein examinees speak their thoughts while responding to assessment items, or eye-tracking studies. These studies are often followed by structured interviews to further probe examinees' cognitive processes. The goals of cognitive lab studies are typically twofold: to improve item accessibility by identifying construct-irrelevant barriers to responding correctly (e.g., points of confusion) and to evaluate whether items elicit cognitive processes consistent with the construct and depth of knowledge intended to be measured by the items. When items elicit the intended cognitive processes, this confirms alignment of the items to content standards and supports the validity of score interpretations for intended uses such as measuring educational achievement.

Since 2017, ACT has conducted several cognitive lab studies and follow-up analyses, and in general, the study results have been consistent with the claim that ACT items elicit evidence of the skills they are intended measure.

Evidence collected through think-aloud protocols for ACT English and reading items largely supported two overarching claims: The test items required that students use targeted skills found in the ELA standards to obtain the correct answer, and the items did not involve construct-irrelevant factors. Most English items included in the study required students to use the context of the sentence and whole passage to answer correctly. Additionally, students showed evidence that they applied knowledge of the conventions of grammar and mechanics and of discourse, such as knowing whether a certain transition word "make[s] more sense" or "flows better" (quotes from study participants). For most reading items, students returned to the passage and applied strategies like skimming, underlining, and summarizing to locate and process relevant information that they used to answer questions. The items clearly required students to use passage evidence—a core component of ELA reading standards—which was illustrated by the way that, for most items, students searched the passage for evidence in order to eliminate options and cited details in the passage as rationales for their answer choices.

During the 2019–2020 school year (before the pandemic shutdown of schools), ACT conducted cognitive lab studies related to the enhanced ACT; these included eye-tracking, a think-aloud protocol for reading, surveys, and guided interviews for reading, science, and mathematics. More complete analysis is included in Chapter 3 of the *Design Framework for the ACT Enhancements*. Participants identified as high scoring tended to use more efficient gaze paths (i.e., eye movement patterns) and could clearly articulate why they selected specific answers with references to the passage.

For the mathematics cognitive lab studies, eye-tracking data provided evidence of cognitive processes. For simple procedural questions, such as finding the median of a dataset, all participants who answered correctly showed gaze paths consistent with the skill required (i.e., reordering the data). As expected, high-scoring students exhibited vision paths consistent with one of the optimal solution paths based on the skill map of the question. Additionally, high scorers did not scan the page repeatedly or require multiple rereads of the stem for more difficult items, which was not the case for low and middle scorers. Problems that required complex problem-solving showed significant differences between the high scorers and low and middle scorers, which was consistent with high school and postsecondary instructor evaluations of problem-solving in the ACT® National Curriculum Survey®. Timing for items was consistent with skill identification, with easy items taking less time than medium-difficulty items, which took less time than difficult items.

During the science cognitive lab studies, low and middle scorers were more likely to spend time looking at the wrong graphic, particularly when the information needed was not in the first graphic presented. This was true even when the question stem specified which graphic was relevant to the question. As with reading passages with graphics (on which more below), students required more time and had more return visits in their gaze paths for less-familiar graphics (e.g., multiple-line graphs, phase diagrams, and process diagrams) than for bar graphs and tables. Students answering items correctly generally followed gaze paths indicating the application of skills as described in the content target of the item. Many students spent significant time rereading the stem or response options multiple times, which could have indicated difficulty decoding the task. Students cited familiarity with the overall topic as making a passage easier.

For the reading test, the visual and quantitative information (VQI) passages were able to elicit evidence of targeted skills that require analysis of multiple sources of information, including both text and graphics. This conclusion was supported by silent eye-tracking analysis of interactions with the VQI passages as well as think-aloud interactions. This study was also used to determine test takers' typical reading rate, which was generally 225–250 standard words per minute unless the student was engaging in speeded behaviors (Arthur & Thomas, 2023). This reading rate was consistent with earlier studies by other researchers. Students used more time per item to answer questions, and this led to an increase in time per item for the enhanced ACT.

In 2024, ACT conducted more cognitive labs based on enhancements to the ACT. A think-aloud protocol was used with students ranging from Grades 10 through 12. Each student was presented with items and passages from the English, mathematics, reading, and science sections. A *q*-matrix was created for each item that included the skills needed to successfully complete the item. The matrix also included common maladaptive skills or errors. As students completed the think-aloud, a proctor coded for the presence of the targeted skills. All these sessions were recorded and coded by a second coder. ACT items are coded only to the highest-order cognitive skill (apex skill), so it is possible for students to use some but not all of the targeted skills and still select one of the distractors. If students are using the targeted skills on an item, they should answer the items correctly at higher rates than students who use one or more maladaptive skills. The results of the cognitive lab showed that the test items, including new item formats, were able to elicit the targeted skills and that students who used the targeted skills answered correctly at higher rates. Across all parts of the ACT, students who used one or more of the targeted skills answered correctly much more frequently than did students who used maladaptive skills. As Table 7.1 shows, students who used only targeted skills answered items correctly over 80% of the time in both the ELA and STEM domains.

**Table 7.1.** Use of Targeted Skills Across Domains

| Domain | Used one or more targeted skills, answered correctly | Used maladaptive skill, answered correctly | Difference between targeted skill and maladaptive skills | Percentage answered correctly when using only targeted skills | Difference between only targeted and maladaptive skills | Overall average correct |
|---|---|---|---|---|---|---|
| **English** | 70.3% | 13.2% | 57.1% | 88.0% | 74.8% | 67.5% |
| **Mathematics** | 73.2% | 19.8% | 53.4% | 84.0% | 64.2% | 54.8% |
| **Reading** | 70.4% | 18.6% | 51.8% | 81.0% | 62.3% | 68.0% |
| **Science** | 73.4% | 21.8% | 51.7% | 94.6% | 72.8% | 67.7% |
| **ELA** | 70.4% | 15.9% | 54.4% | 84.5% | 68.6% | 67.8% |
| **STEM** | 73.3% | 20.8% | 52.6% | 89.3% | 68.5% | 61.2% |

To determine whether changes to item formats (such as the addition of stems to English items) were minimizing construct-irrelevant variance, ACT asked participants whether they understood what the question was asking them to do, even if they did not know how to correctly answer it (see Table 7.2). We also asked them whether they were able to show what they knew and could do on each section of the test (refer to Table 7.3). For all subjects, more than 85% of the students responded that they understood what the question was asking them to do. Over 70% agreed that they were able to show what they knew and could do on each section of the test (Table 7.3).

**Table 7.2.** Student Ability to Understand What the Question Was Asking by Test Section

I was able to understand what the question was asking me to do, even if I did not know how to answer the specific question on the ____ test.

| Section | % Strongly disagree | % Disagree | % Agree | % Strongly agree |
|---|---|---|---|---|
| **English** | 0.00 | 0.00 | 44.44 | 55.56 |
| **Mathematics** | 0.00 | 14.81 | 44.44 | 40.74 |
| **Reading** | 0.00 | 3.70 | 51.85 | 44.44 |
| **Science** | 3.70 | 7.41 | 48.15 | 40.74 |

**Table 7.3.** Students' Rating of Their Ability to Show What They Know and Can Do by Test Section

I was able to show what I know and can do in the subject on the ____ test.

| Section | % Strongly disagree | % Disagree | % Agree | % Strongly agree |
|---|---|---|---|---|
| **English** | 0.00 | 11.11 | 25.93 | 62.96 |
| **Mathematics** | 3.70 | 25.93 | 33.33 | 37.04 |
| **Reading** | 0.00 | 11.11 | 29.63 | 59.26 |
| **Science** | 0.00 | 25.93 | 29.63 | 44.44 |

The collection of evidence of cognitive processes strengthens claims about what students know and can do. A more complete discussion of all the evidence of cognitive processes can be found in Chapter 3 of the *Design Framework for the ACT Enhancements* (Allen et al., 2026).

### *7.1.3 Statistical Relationships Between High School Coursework and Grades and ACT Scores*

ACT tests are oriented toward the general content areas of high school and college curricula. In particular, ACT conducts a National Curriculum Survey every 3 to 5 years to ensure that our assessments always measure the knowledge and skills being taught in schools around the country (ACT, 2020b; see Chapter 2 of this manual). Students' performance on the ACT should therefore be related to the high school courses they have taken and to their performance in these courses.

One component of registering for the ACT is completing the Course/Grade Information Section (CGIS), which collects information on about 30 high school courses in English, math, social studies, natural sciences, languages, and arts. Many of these courses form the basis of a high school college-preparatory curriculum and are frequently required for college admission or placement. For each of the 30 courses, students indicate whether they have taken or are currently taking the course, whether they plan to take it, or whether they do not plan to take it. If they have taken the course, they indicate the grade they received (A–F). Self-reported coursework and grades collected with the CGIS have been found to be accurate relative to information provided on student transcripts (Sanchez & Buddin, 2016; Sawyer et al., 1988; Valiga, 1986; see also the next section).

Table 7.4 displays the ACT scale score means and the percentage of students meeting the College Readiness Benchmarks for the English, mathematics, reading, and science tests by common high school course-taking patterns in English, math, social studies, and science. As seen in the table, mean test scores and the percentage of students meeting the Benchmarks typically increase as the number of years of coursework increases in that subject area. However, the strength (and sometimes direction) of this relationship is also conditional on the types of courses taken. For example, students who take math courses through calculus and science courses through physics typically have higher mean scores and Benchmark attainment rates than students who complete a different sequence of courses over the same number of years of coursework.

**Table 7.4.** College Readiness Benchmark Attainment Percentages and Average ACT Scores by Common Course Patterns, ACT-Tested High School Graduating Class of 2021

| Course pattern | | *n* | Mean ACT score | Percentage (%) meeting Benchmark |
|---|---|---|---|---|
| **English** | Eng. 9, Eng. 10, Eng. 11, Eng. 12, other | 165,958 | 22.3 | 72 |
| | Eng. 9, Eng. 10, Eng. 11, Eng. 12 | 623,323 | 21.1 | 65 |
| | Less than 4 years of English | 56,980 | 16.6 | 38 |
| | Zero years / no English | 449,088 | 16.8 | 39 |
| **Mathematics** | Alg. 1, Alg. 2, geom., trig., calc. | 41,584 | 23.2 | 61 |
| | Alg. 1, Alg. 2, geom., trig, other adv. math | 58,940 | 21.4 | 49 |
| | Alg. 1, Alg. 2, geom., trig. | 38,353 | 18.8 | 26 |
| | Alg. 1, Alg. 2, geom., other adv. math | 179,211 | 19.1 | 28 |
| | Other comb. of 4 or more years of math | 326,953 | 24.1 | 66 |
| | Alg. 1, Alg. 2, geom. | 98,409 | 16.5 | 9 |
| | Other comb. of 3 or 3.5 years of math | 50,170 | 19.5 | 32 |
| | Less than 3 years of math | 44,627 | 16.0 | 9 |
| | Zero years / no math | 457,102 | 18.1 | 22 |
| **Social studies** | U.S. hist., world hist., Am. gov., other hist. | 25,929 | 22.6 | 53 |
| | Other comb. of 4 or more years soc. sci. | 400,925 | 23.3 | 57 |
| | U.S. hist., world hist., Am. gov. | 63,786 | 19.9 | 37 |
| | Other comb. of 3 or 3.5 years of soc. sci. | 247,455 | 22.1 | 50 |
| | Less than 3 years of social sciences | 99,996 | 19.8 | 36 |
| | Zero years / no social sciences | 457,258 | 18.5 | 30 |
| **Science** | Gen. sci., bio., chem., phys. | 326,363 | 22.9 | 52 |
| | Bio., chem., phys. | 109,874 | 24.0 | 60 |
| | Gen. sci., bio., chem. | 244,298 | 20.4 | 33 |
| | Other comb. of 3 years of nat. sci. | 25,088 | 19.2 | 26 |
| | Less than 3 years of natural science | 131,559 | 18.5 | 22 |
| | Zero years / no natural science | 458,167 | 18.5 | 23 |

Moreover, as shown in Table 7.5, students who have completed or plan to complete a core curriculum tend to achieve higher ACT scores than those who have not completed a core curriculum (ACT, 2016b), where a core curriculum is defined as at least 4 years of English and at least 3 years each of math, social studies, and natural sciences. From 2011–2012 through 2015–2016, the ACT Composite scores of students who completed a core curriculum averaged about 3 scale score points higher than the scores of those who did not.

**Table 7.5.** Average ACT Scores by Academic Preparation, 2012–2016

| Academic preparation | Year | *n* | ACT score | | | | |
|---|---|---|---|---|---|---|---|
| | | | English | Math | Reading | Science | Composite |
| Core curriculum or more completed | 2011–12 | 1,259,744 | 21.3 | 21.8 | 22.0 | 21.6 | 21.8 |
| | 2012–13 | 1,322,739 | 21.2 | 21.7 | 22.0 | 21.5 | 21.7 |
| | 2013–14 | 1,347,997 | 21.4 | 21.7 | 22.2 | 21.6 | 21.8 |
| | 2014–15 | 1,389,338 | 21.4 | 21.7 | 22.3 | 21.8 | 21.9 |
| | 2015–16 | 1,441,538 | 21.3 | 21.5 | 22.3 | 21.7 | 21.9 |
| Core curriculum not completed | 2011–12 | 355,849 | 18.3 | 19.1 | 19.4 | 19.1 | 19.1 |
| | 2012–13 | 396,592 | 17.8 | 18.9 | 19.0 | 18.8 | 18.7 |
| | 2013–14 | 405,073 | 17.9 | 18.9 | 19.2 | 18.9 | 18.9 |
| | 2014–15 | 424,562 | 18.0 | 18.9 | 19.3 | 19.0 | 18.9 |
| | 2015–16 | 483,335 | 17.8 | 18.7 | 19.2 | 18.8 | 18.7 |

The findings shown in Tables 7.4 and 7.5 are consistent with the notion that the ACT is a curriculum-based test. Additionally, an analysis by McNeish et al. (2015) showed that, in general, coursework and high school grades were strongly associated with performance on the ACT after specific other factors were statistically controlled for in the model. However, it is also conceivable that other factors not accounted for by the model, such as students' noncognitive characteristics, account for the observed association between high school coursework and ACT scores. In the McNeish et al. study, the researchers investigated the relationships between noncognitive characteristics, high school coursework and grades, school characteristics, and ACT test scores. They found that between 44% and 61% of the variance in ACT scores was explained by high school grade point average (HSGPA), coursework taken, school characteristics, noncognitive characteristics, and demographic characteristics. High school academic factors, such as HSGPA, courses taken, and advanced coursework, accounted for the most variance in all five ACT scores ($R^2$ = 0.28 to 0.46). These three factors accounted for 64% to 77% of the total variance explained by the models. In particular, taking higher-level mathematics and science courses and subject-relevant accelerated, advanced, honors, or dual-enrollment courses was associated with sizable mean ACT score differences. Specific English and social studies courses were not included in the models because of the limited variability in students' course-taking in these subject areas and their collinearity with other variables, such as coursework taken in math and science. The findings from this study are consistent with those from earlier studies that examined relationships among coursework, grades, and ACT scores (Noble et al., 1999a, 1999b; Noble & McNabb, 1989; Schiel et al., 1996).

ACT research has shown that taking rigorous college-preparatory math courses is associated with higher ACT mathematics and Composite scores (e.g., ACT, 2016b; Noble et al., 2001, 2006). Schiel et al. (1996) statistically controlled for prior achievement using ACT Plan scores and found substantive increases in average ACT mathematics and science scores associated with taking higher-level math and science courses. In other studies, researchers found that, in a typical high school, students who take higher-level math or science courses (e.g., trigonometry, calculus, chemistry, or physics) can expect to earn meaningfully higher average ACT mathematics and science scores than students who do not take such courses (ACT, 2005; Noble & Schnelker, 2007). However, the expected benefits of coursework taken in high school for increasing ACT performance depend on the student's high school, regardless of prior achievement and grade level at testing (Noble & Schnelker, 2007).

Extending this line of research to a current national sample, Sanchez (2025b) examined how rigorous high school coursework relates to ACT Composite scores using data from 453,439 students in the 2023 ACT-tested graduating class. Using hierarchical linear models with students nested within schools, the study found that completing math coursework beyond Algebra 2 and science coursework through physics was strongly associated with higher ACT Composite scores, even after the study controlled for student and school characteristics. Score gains associated with advanced coursework were larger for White, Asian, and higher-income students than for students from historically underserved groups; these results underscore the importance of equitable access to rigorous coursework as a pathway to college readiness.

### *7.1.4 Statistical Relationships Between High School Coursework and Grades and ACT Benchmark Attainment*

To provide students and educators with an empirical definition of what it means to be academically ready for first-year credit-bearing college courses, ACT developed the ACT College Readiness Benchmarks based on college course grade data from 214 two- and four-year institutions (Allen, 2013). The ACT College Readiness Benchmarks are scores on the ACT multiple-choice tests that represent the level of achievement required for students to have at least a 50% chance of obtaining a B or higher in related first-year college courses. The Benchmarks also correspond to an approximate 75% chance of earning a C or higher in these courses. The Benchmarks link to common first-year courses: ACT English to English Composition I, ACT mathematics to college algebra, ACT reading to social science courses, and ACT science to biology. The Benchmarks correspond to scores of 18, 22, 22, and 23 on the English, mathematics, reading, and science tests, respectively. For more details on the development of the Benchmarks, as well as details on the ACT STEM and ELA Benchmarks, see Chapter 5.

A study by Ling and Radunzel (2017) examined how high school coursework taken and grades earned related to students' chances of meeting the ACT College Readiness Benchmarks in each of the four test sections after accounting for other student and school characteristics. The study findings indicated that students who took rigorous courses in high school and earned good grades were more likely to meet the Benchmarks and therefore more likely to succeed in first-year college courses. These study findings are consistent with those from an earlier ACT study by Noble and Schnelker (2007), which indicated that some courses and course sequences were more strongly associated with Benchmark attainment than others. Each incremental college-preparatory course taken, particularly in math and science (e.g., trigonometry beyond Algebra 2, physics beyond chemistry), increased the likelihood of meeting the Benchmarks more than the number of courses taken in a discipline alone. A limitation of these studies is that students' self-reported courses taken and grades earned are based only on those courses available in the ACT CGIS, which does not provide more detailed information on courses taken, especially in English.

Building on earlier Benchmark research, Sanchez (2025c) examined how high school coursework and grades relate to attainment of the ACT ELA and STEM Benchmarks using data from 176,095 students in the 2024 ACT-tested graduating class. Using logistic regression with cluster-robust standard errors, the study found that STEM Benchmark attainment was most strongly associated with math GPA, taking math courses beyond Algebra 2, completing a biology-chemistry-physics course sequence, and having taken advanced math coursework. Math GPA was a considerably stronger predictor of STEM Benchmark attainment than science GPA. For ELA Benchmark attainment, the strongest predictors were English GPA, having taken advanced English coursework, and having taken advanced social studies coursework. These results corroborate earlier single-subject Benchmark findings and provide the first large-scale evidence for how coursework and grades relate to the ELA and STEM Benchmarks specifically.

### *7.1.5 Distinction Between ACT Scores and HSGPA as Measures of Educational Achievement*

ACT scores are statistically associated with high school grades (Table 7.6; see also the previous section). That is, students who have higher HSGPAs tend to achieve higher ACT scores. As shown in Table 7.6, students in the highest HSGPA category (3.50–4.00) tend to earn ACT Composite scores nearly 10 scale score points higher than students in the lowest HSGPA category (0.00–1.99). However, ACT scores and HSGPAs are different measures in that there are some noncognitive predictors related to high school grades that are not directly related to ACT scores (McNeish et al., 2015; Noble et al., 1999a, 1999b). To the extent that grades measure educational achievement, there will be a strong statistical relationship between grades and ACT scores. However, grades are more subjective than standardized test scores because of differing grading standards and practices (Brookhart, 1993; Pilcher, 1994; Stiggins et al., 1989). Within a given school, teachers may differ in the criteria they use to judge student achievement. Effort and reward are often confounded with academic accomplishment in assigning course grades (Allen, 2005; Pilcher, 1994; Willingham et al., 2002). In a review of the literature on elementary and high school grading practices over the past century, Brookhart (2015) concluded that “report card grades can be reliable and valid measures of graded achievement, but may not be depending on individual teachers’ grading practices” (p. 268). Grading practices also vary across schools; an A in one school may be equivalent to a C in another (United States Department of Education, 1994). Consequently, the interpretation of high school grades should take into account differences across high school curricula and grading standards. Grade inflation also adversely affects the usefulness of high school grades in terms of understanding educational achievement.

**Table 7.6.** Average ACT Score by HSGPA Ranges, 2015–2016

| HSGPA group | Number | English | | Math | | Reading | | Science | | Composite | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Mean | *SD* | Mean | *SD* | Mean | *SD* | Mean | *SD* | Mean | *SD* |
| 3.50–4.00 | 726,643 | 24.5 | 6.0 | 24.2 | 5.2 | 25.1 | 5.9 | 24.2 | 5.1 | 24.6 | 5.0 |
| 3.00–3.49 | 479,292 | 19.5 | 5.5 | 19.8 | 4.4 | 20.7 | 5.5 | 20.3 | 4.5 | 20.2 | 4.4 |
| 2.50–2.99 | 274,467 | 16.9 | 5.0 | 17.7 | 3.6 | 18.4 | 5.1 | 18.3 | 4.3 | 18.0 | 3.9 |
| 2.00–2.49 | 154,002 | 15.1 | 4.6 | 16.5 | 3.0 | 16.8 | 4.7 | 16.8 | 4.1 | 16.4 | 3.5 |
| 0.00–1.99 | 75,255 | 13.6 | 4.3 | 15.7 | 2.5 | 15.4 | 4.2 | 15.5 | 3.9 | 15.2 | 3.1 |
| All | 2,090,342 | 20.1 | 6.8 | 20.6 | 5.4 | 21.3 | 6.5 | 20.8 | 5.6 | 20.8 | 5.6 |

### Accuracy of Self-Reported Coursework and Grades

The accuracy of the high school course and grade information students provide in the ACT registration folder within the CGIS is a focus of continuing research at ACT. Sanchez and Buddin (2016) concluded that students' self-reported grade information accurately represented students' high school experience. About 94% of students accurately reported taking particular courses. The correlation between self-reported and transcript course grades was .66, with 96% of self-reported grades within a single letter grade of the transcript grade. HSGPA computed from self-reported course grades was highly correlated with transcript grade point average ($r$ = .83). The accuracy of coursework and grades differed little by gender, race/ethnicity, or income. The results indicated that self-reported coursework and grades are reasonably accurate measures for use in education research and for preliminary screening by college admissions officials.

### Differential Grading Standards

A study by Woodruff and Ziomek (2004b) assessed how grading standards varied across high schools. This study found that grades are more of a relative (as compared to an absolute) standard in that the interpretation of grades can vary from school to school. That is, an A at one high school does not necessarily translate to the same level of academic achievement as an A at another school. In addition to analyzing overall HSGPA, this study evaluated differential grading standards by subject area. For further details, see the full ACT research report (Woodruff & Ziomek, 2004b).

Grade inflation and differential grading standards introduce additional variability into high school grades, allowing them to differ in value from year to year and school to school. In contrast, the ACT is carefully constructed to measure the same content and have the same statistical properties from year to year, and its administration does not vary from school to school. Hence, ACT scores are a useful supplement to high school grades for those attempting to make valid evaluations of college readiness.

## *7.1.6 Statistical Relationships Between End-of-Course Exams and ACT Scores*

If performance on the ACT test is influenced by mastery of high school course content, one would expect that standardized measures of achievement in specific high school courses would predict performance on the ACT. Moreover, the predictive relationship should be apparent even when studies control for students' levels of achievement before high school. To test this proposition, a study examined the extent to which ACT scores are predicted by scores from end-of-course exams when pre–high school academic achievement is controlled for (Allen, 2015).

The results of the study support the proposition that performance on the ACT is related to standardized measures of achievement in high school courses in the core subject areas (English, math, social studies, and natural science). Thus, the study results can be used as a source of evidence for validating the use of ACT scores as measures of educational achievement. The predictive weights (standardized regression coefficients) of the end-of-course exams with closer time proximity to the ACT were larger than the predictive weight of the pre–high school achievement measure (ACT Explore) from the same subject area. While ACT Explore scores were strong predictors of ACT scores, results showed that achievement in core high school courses also had a strong relationship with ACT scores. That is, students who mastered content from core high school courses were more likely to achieve high ACT scores. In comparison to the McNeish et al. (2015) study (discussed in Section 7.1.3), the models in this study explained a greater percentage of the variation in ACT scores. ACT Explore scores predicted ACT scores better than high school course grades and courses taken, high school characteristics, noncognitive characteristics, socioeconomic status, and demographic variables. This was likely due to the Explore scores and the end-of-course exam scores being more directly related to the outcome, which was also a standardized measure of academic achievement, than to unstandardized variables such as high school coursework and grades.

In addition to evidence indicating that end-of-course exam performance predicts ACT scores, a recent study suggests that ACT test scores can also be used to help identify academically prepared students who might benefit from more rigorous high school courses, including Advanced Placement (AP) courses (Radunzel & Allen, 2020). The study found that ACT test scores were positively related to AP exam scores and were good predictors of success. The links to AP exam success were defined in two ways (receiving a score of 3 or higher on the AP exam or receiving a score of 4 or higher on the AP exam) and were developed in relation to content-relevant scores for most courses.

### *7.1.7 Statistical Relationships Between Noncognitive Factors and ACT Scores*

ACT has conducted substantial research examining the relationship between ACT scores and noncognitive student characteristics such as educational plans, academic behaviors, and perceptions of self. Some of these characteristics should be considered aspects of social and emotional learning, which encompasses a broad array of interpersonal, self-regulatory, and task-related behaviors important for adaptation to and successful performance in education and workplace settings (Casillas et al., 2015).

When students register for the ACT, they are asked to provide information about their background, interests, needs, and plans in the Student Profile Section (SPS) of the registration form. Correlations were calculated between selected variables and ACT scores for the 2016

ACT-tested graduating class. As shown in Table 7.7, students who described their high school curriculum as college-preparatory in nature ($r$ = .31 to .35) and aspired to higher educational levels ($r$ = .32 to .36) tended to have higher ACT scores. Those who reported not needing help with their reading ($r$ = .02 to .17), study skills ($r$ = .07 to .09), or math skills ($r$ = .10 to .26) also tended to have higher ACT scores.

**Table 7.7.** Correlations Among ACT Scores and Background Characteristics

| ACT score | College-preparatory curriculum [a] | Educational plans [b] | Does not need help in [c] | | |
|---|---|---|---|---|---|
| | | | Reading | Study skills | Math |
| English | .34 | .34 | .14 | .08 | .13 |
| Mathematics | .32 | .33 | .02 | .08 | .26 |
| Reading | .31 | .32 | .17 | .07 | .10 |
| Science | .31 | .32 | .09 | .08 | .18 |
| Composite | .35 | .36 | .12 | .09 | .18 |

*Note.* All $p$-values < .0001.

[a] Responses were coded 1 (college preparatory) and 0 (business or commercial, vocational-occupational, other, or general).

[b] Responses were coded 1 to 5 (vocational-technical program, associate's degree, bachelor's degree, 1 to 2 years of graduate program, professional degree).

[c] Responses were coded 1 (do not need assistance) and 0 (need assistance).

A study by McNeish et al. (2015) examined the relationships between students' noncognitive characteristics, high school coursework and grades, school characteristics, demographic characteristics, and ACT scores, with an emphasis on noncognitive measures related to students' academic goals, behaviors, perceptions, and parental involvement. Consistent with earlier studies by Noble et al. (1999a, 1999b), this study's results indicated that noncognitive variables provided incremental improvement to the prediction of ACT scores; however, relative to coursework taken, grades, and school characteristics, the effect was small (less than 10%). It is important to note that students' noncognitive characteristics alone explained 29% of the variance in HSGPA, highlighting the importance of noncognitive factors to future academic achievement (for further details, see McNeish et al., 2015). Given that HSGPA entered the model first, any overlap in variance accounted for in ACT scores by HSGPA and noncognitive characteristics would be attributed to HSGPA.

In a study by Noble et al. (2006), structural equation modeling results indicated that ACT scores were directly related only to academic achievement in high school as measured by grades earned and coursework taken. Education-related accomplishments, as well as activities and perceptions of self and others (noncognitive measures), had only indirect relationships to ACT scores through academic achievement.

In sum, findings from these studies suggest that noncognitive characteristics are associated with students' choices of high school coursework and the grades they earn in those courses, which, in turn, are strongly related to ACT scores.

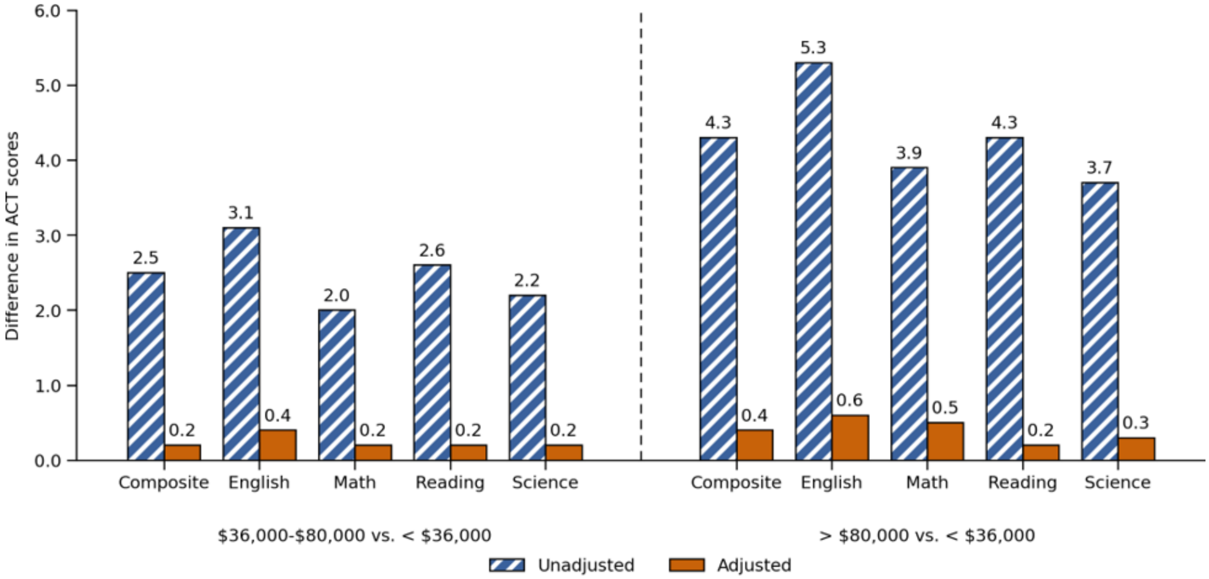
Social and emotional skills such as sustaining effort (how your actions demonstrate diligence, effort, organization, self-control, and compliance with rules) or maintaining composure (how your actions demonstrate relative calmness, serenity, and the ability to manage emotions effectively) are related to many positive outcomes, including academic achievement (e.g., Poropat, 2009).

### *7.1.8 Understanding Group Differences on the ACT*

Equity and fairness are important concerns for educators. Researchers have examined the strength of the associations between ACT performance and predictors such as coursework, course grades, student and high school characteristics, and educational plans by race/ethnicity, gender, and annual family income (e.g., Chambers, 1988; Noble et al., 1992; Noble et al., 1999a, 1999b; Noble & McNabb, 1989). Their findings suggest that differential performance is largely attributable to differential academic preparation across student demographic groups.

Results from a study by McNeish et al. (2015) support the hypothesis that differential performance on the ACT results from differential academic preparation, regardless of race/ethnicity, gender, or annual family income. Specifically, after accounting for HSGPA, high school coursework, school characteristics, and other noncognitive factors, the study found that socioeconomic status and other demographic characteristics collectively accounted for only a small percentage of the variance in ACT scores (4% or less). Additionally, differences in ACT scores among racial/ethnic, family income, and parental education level groups were substantially reduced when students' academic preparation levels were taken into account. School-level demographic characteristics, along with other school-level characteristics, were included in the models to account for high school attended. For example, students from high-income households (>$80,000) tended to earn ACT Composite scores that were 4.3 points higher than students from low-income households (<$36,000; see Figure 7.1). However, after the researchers accounted for student and school-level differences among the two groups, the adjusted difference in ACT Composite score was reduced to 0.4.

**Figure 7.1.** Unadjusted and Adjusted Mean Differences in ACT Scores by Family Income




Findings from the McNeish et al. (2015) study are consistent with results from earlier studies on this topic (Noble et al., 1999b; Schiel et al., 1996). Moreover, a more recent study that focused on deriving a high school academic rigor index (HSAR) revealed that differences in the HSAR helped explain differences in ACT Composite scores across racial/ethnic groups (Allen et al., 2019). The HSAR was derived by optimizing the prediction of first-year college grades based on high school courses taken, grades, and indicators of advanced coursework. In the study results, the mean difference in ACT Composite score for White and African American students was reduced from 4.4 points to 2.8 points after the researchers adjusted for the HSAR index.

### *7.1.9 Relationship Between Test Preparation and ACT Performance*

The ACT assessment measures much of the knowledge and many of the skills taught in high school. Thus, it would stand to reason that long-term learning in school, rather than short-term preparation focused on test format or test-taking skills, would be the best form of preparation for the ACT. To better understand the relationship between test preparation and ACT scores, several studies examined score increases associated with short-term test preparation activities. An analysis by Scholes and Lain (1997) suggested that preparing with practice tests for two or more hours was associated with slightly higher ACT Composite scores for first-time testers when HSGPA and grade level were controlled for, but using workbooks or taking a preparation course were not. In a follow-up study, Scholes and McCoy (1998) estimated the average difference in ACT Composite scores between examinees who did and did not participate in different types of short-term preparation (workbooks and courses, workshops, or computer software) and long-term preparation (taking or planning to take a recommended core curriculum and taking or planning to take advanced courses in math or science). Results showed that long-term preparation was related to much higher scores for first-time testers than short-term strategies.

In a series of three studies, Schiel and Valiga (2014a, 2014b, 2014c) investigated the association between test preparation activities and ACT performance for repeat testers. In these studies, repeat testers were surveyed about their engagement in test preparation before taking the ACT test. The results revealed that average ACT Composite scores increased slightly from the first to the second test regardless of preparation. The observed increase, even for students who did not engage in supplemental preparation, possibly reflects a practice effect from taking the test previously or the effect of additional classroom instruction between testing occasions (Camara & Allen, 2017). In general, students who engaged in test preparation activities before taking the ACT for the second time exhibited a slightly greater average ACT Composite score gain compared to those who did not. A similar result was found in two follow-up studies that moved beyond descriptive statistics by statistically controlling for student background characteristics (e.g., race/ethnicity, gender, first ACT Composite score) and test-related behaviors (e.g., experiencing notable stress or anxiety, having adequate sleep). First, a study by Schiel (2020) that used a similar data source from a more recent cohort of students found that the adjusted average ACT Composite score gains were larger for students who prepared for the second test than for those who did not (1.22 points versus 0.85 points, respectively). A subsequent study employed quasi-experimental methods to allow for a causal examination of the efficacy of test preparation (Moore et al., 2019). Results indicated that students who prepared for a second ACT test earned Composite scores that were 0.71 points higher on average than the scores of students who did not prepare.

Most survey respondents from the Schiel and Valiga studies (2014a, 2014b, 2014c) indicated that test preparation built their confidence, familiarized them with the test, refreshed their memory of the content areas, and helped them understand subject matter. Respondents who found test preparation useful for their second test exhibited greater average score increases than respondents who did not.

Further analyses from the Schiel and Valiga studies (2014a, 2014b, 2014c) revealed that test takers who spent more time engaging in test preparation activities tended to increase their scores more than those who spent less time. Other studies that have focused on the use of specific test preparation products have reached similar conclusions. For example, two studies by Sanchez (2019, 2020) found a positive association between greater ACT Online Prep (AOP) usage and ACT score gains, with AOP usage defined by variables such as the number of active days in AOP, the number of practice sessions completed, the number of full-length practice tests completed, the number of system resets, and the number of hours spent on AOP activities. Sanchez (2020) reported that students who used AOP for less than 7 hours saw greater ACT Composite score gains on average than students who did not use test preparation (1.06 versus 0.67). In comparison, the average score gains for students who used AOP for 7 or more hours were more than double those for students who did not use test preparation (1.29–1.37 versus 0.67). Similarly, Payne and Allen (2019) found that ACT Composite score gains increased with more time spent on ACT Academy—a now-retired free online platform that provided assessments and individualized learning plans to help students improve their academic skills. For example, compared to the score gains of students who did not engage in test preparation before retesting, the average ACT Composite score gain was 0.5 points higher for students who spent 1 hour using ACT Academy and 1.3 points higher for students who spent at least 6 hours using ACT Academy.

In these studies, average increases were approximately 1–2 points depending on the duration of preparation. These studies suggest that the average effect of preparation is small but that larger score increases (and decreases) should be expected for individual students and for different student subgroups. For example, Sanchez (2018) found that, while students from all racial/ethnic groups benefitted from AOP, the average benefit to students of color (about 1 point or more) was greater than the average benefit to White students (under half a point). Results from the various studies described suggest some reasons for the relationship between test preparation and higher ACT performance, such as increasing confidence, becoming familiar with the test format, and refreshing or learning subject matter.

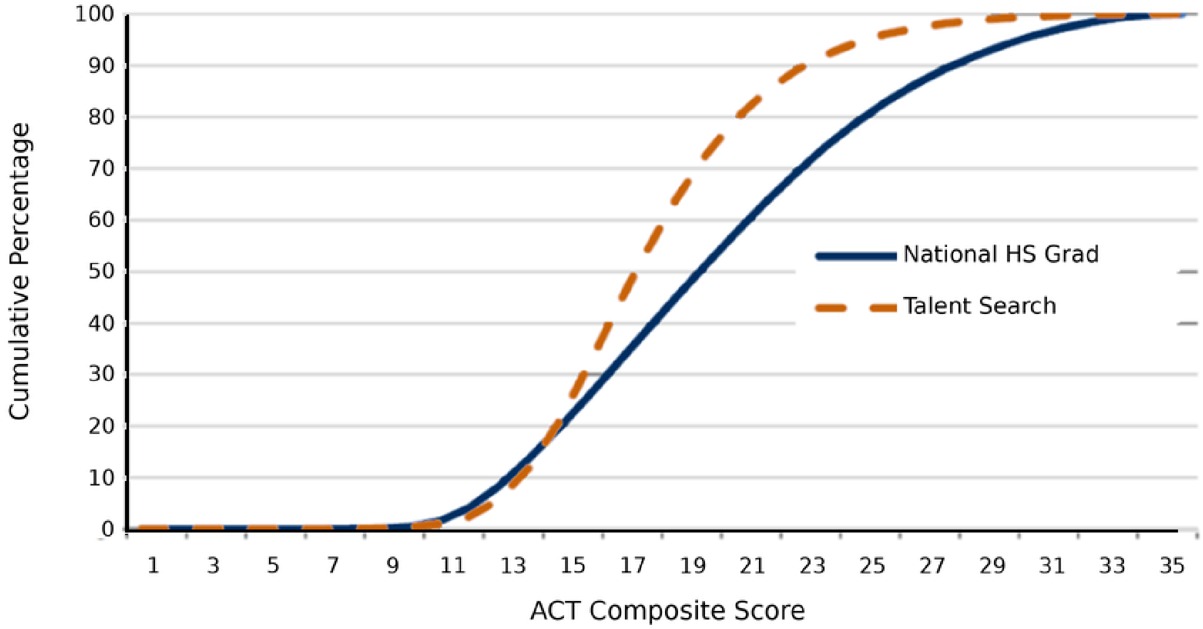
### *7.1.10 Using ACT Scores to Identify Students With Very High Academic Achievement for Gifted and Talented Programs*

ACT scores have, over the years, been used successfully by national talent search programs to identify academically talented young people. Talent search programs provide these individuals with such services as advanced-level course materials, recognition ceremonies, and special residential programs. In a typical talent search program, students in Grade 7 or 8 who score very high (e.g., top 3%) on in-school standardized achievement tests are invited by the program to take the ACT. Those applicants earning very high ACT scores are then invited to participate in a special residential or recognition program.

Figure 7.2 displays two ACT Composite score cumulative distributions, one representing the scores of 2016 high school graduates and the other representing the scores of a group of talent search applicants. The score distribution for the 2016 high school graduates ($N$ = 2,090,342) in this figure was based on students who took the ACT during their sophomore, junior, or senior year and who graduated from high school in the spring of 2016. Only the most recent ACT score of each high school student was retained for analysis. The score distribution for talent search applicants was based on data from 40,562 students who took the ACT during Grade 6, 7, or 8 in 2016 and sent their scores to a particular national talent search program.

Figure 7.2 reveals that the cumulative distribution for the 2016 ACT-tested graduating class is shifted slightly to the right of the cumulative distribution for the students identified by the talent search program (average ACT Composite score of 20.8 versus 18.6, respectively). This figure suggests neither a floor effect (i.e., the assessment is too challenging for this student population and everyone scores near the bottom of the score scale) nor a ceiling effect (i.e., the assessment is too easy for this student population and everyone scores near the top of the score scale), indicating that ACT scores can meaningfully assess the educational achievement of academically talented students in Grades 6–8. This is further substantiated by a study that compared ACT Composite scores earned by 7th-graders to ACT Composite scores earned by the same students 4 or 5 years later (Allen, 2016a). In that study, the average score gain was 9.4 points. This suggests a lack of ceiling on the ACT for such students, though average score gains were lowest for students who scored near the top of the ACT Composite scale as 7th-graders.

**Figure 7.2.** ACT Composite Cumulative Percentages for 2016 ACT-Tested High School Graduates and Talent Search 6th-, 7th-, and 8th-Grade Students




A study by Schiel (1998) examined the academic benefits in high school of an intensive summer program for academically talented 7th-graders. The results of the study suggested that participation in summer residential programs is positively related to academically talented students' subsequent academic performance in high school. For more details, see the full ACT research report (Schiel, 1998).

Allen (2025) developed cumulative percentile rank tables for ACT scores that were earned by examinees in Grades 7, 8, and 9. These populations are typical of those identified by talent programs. The norming sample consisted of 49,446 ACT records (2,420 in Grade 7; 6,375 in Grade 8; 40,651 in Grade 9) administered between September 2022 and July 2025. The mean ACT Composite scores (without science) were 18.6 (*SD* = 5.1) for Grade 7, 19.9 (*SD* = 5.5) for Grade 8, and 19.1 (*SD* = 5.6) for Grade 9; these means exceed PreACT 8/9 norms (15.9 fall, 16.6 spring) and PreACT 9 norms (16.4 fall, 17.0 spring), which is consistent with a higher-achieving talent-identification population. Earlier research (Allen, 2016a) suggested that students who take the ACT for the first time in Grade 7 gain an average of 9.4 scale score points by the time they retest in Grade 11 or 12. These norms support this interpretation for talent identification examinees with the understanding that they are not nationally representative.

### 7.1.11 Score Interpretations for Examinees Using English Learner Supports

As described in Chapter 4, some English learners (ELs) are eligible for language supports while taking the ACT. Analyses from a preliminary study suggest that providing supports and accommodations on the ACT to ELs may improve their scores and result in a more accurate reflection of their true achievement levels (Moore et al., 2018). Data for the study were gathered in states that administered the ACT statewide, captured students' EL status, and provided state-approved accommodations to EL students.

More recently, based on an analysis of data including ELs who took the ACT using ACT-approved supports, a study indicated that the reliability of ELs' ACT scores was comparable to that of their scores on other assessments and sufficiently high that it did not, by itself, raise concerns about the validity of the scores (Moore et al., 2020). Data for the study included a national sample of students who took the ACT test as part of 2018 state and district testing (10,235 EL and 26,378 non-EL students). ELs who tested with (27%) and without (73%) ACT-approved supports were included in the study. In the study, classification consistency analyses revealed similar agreement rates for ELs and non-ELs, and there was no evidence of differential item functioning for ELs. The latter finding suggests that item characteristics did not introduce additional bias that would have raised concerns about score validity for ELs. Prior studies on EL students revealed a slight underprediction of first-year college outcomes based on college admissions test scores. It is hypothesized that the use of appropriate accommodations by ELs (such as those available for the ACT test) results in ACT scores that are better predictors of first-year college outcomes because these supports reduce construct-irrelevant variance due to non-proficiency in English. Empirical studies to substantiate this hypothesis are currently planned.

Moore and Schnieders (2026c) and Moore et al. (2026) explored the relationships between WIDA ACCESS English learning proficiency (ELP) scores and Grade 11 ACT scores for students in two states. State 1 included 642 ELs (182 of whom tested with ACT supports) and 35,234 non-ELs from the 2017–2018 through 2023–2024 academic years. State 2 included 15,788 ELs (3,225 of whom tested with ACT supports) and 143,461 non-ELs from the 2018–2019 through 2022–2023 academic years. In both states, EL students scored about 5 to 6 ACT Composite points lower than non-ELs. The correlations between overall ACCESS scores and ACT Composite scores were moderate to strong, ranging from $r = .48$ in State 1 to $r = .42$ in State 2. Correlations in subdomains were varied. ACCESS Reading correlations with the ACT Composite ranged from $r = .45$ to $r = .53$, while ACCESS Speaking and Writing had weaker correlations, ranging from $r = .22$ to $r = .27$ for Speaking and from $r = .25$ to $r = .30$ for Writing. In both states, ACT scores rose with ACCESS Reading level. Composite means at ACCESS Reading Level 1 were 12.4 (State 1) and 12.2 (State 2), rising to 15.7 (State 1) and 15.3 (State 2) at Level 6. Moore and Schnieders (2026d) looked at the findings from both states and concluded that ACT score interpretations for examinees at the lowest ACCESS Reading levels should be made with caution given the potential for construct-irrelevant variance associated with limited English proficiency at those levels.

Moore and Schnieders (2026a) looked into the effect of approved ACT testing supports for EL students in State 2. The supports include time and a half on a single testing day, ACT-provided translated test instructions, and an approved word-to-word bilingual dictionary. After the researchers made adjustments for ACCESS Reading level, free or reduced-price lunch eligibility, Grade 11 GPA, credits earned, and advanced coursework, EL students who tested with supports scored 0.45 points higher than ELs without supports (0.70 points higher on English, 0.18 points higher on mathematics, 0.56 points higher on reading, and 0.36 points higher on science). About 20% of ELs examined in State 2 received supports, which indicates substantial room for wider awareness and participation.

Moore and Schnieders (2026b) showed that in State 2, ELs were less likely than non-ELs to have taken advanced or college-credit coursework (advanced/honors: 34% versus 51%; college or dual enrollment: 2% versus 10%), less likely to take calculus or other higher math (0.4% versus 5%) or precalculus, trigonometry, or statistics (2% versus 17%), and less likely to take physics (10% versus 25%). Course-taking increased among ELs with increasing ACCESS Reading level. Advanced or honors enrollment was 25% at Level 1 and 48% at Level 6. After Moore et al. (2026) added Grade 11 GPA, credits earned, and course rigor into regression models that already included ACCESS Reading level, EL supports, and demographic covariates, $R^2$ increased by between 0.01 and 0.04 depending on the model (it was smallest for reading and largest for the Composite), and a dominance analysis showed that course rigor and credits combined explain 10% to 20% of the total adjusted $R^2$ in predicting ACT scores among EL students.

### *7.1.12 Grade Inflation and Its Implications for ACT Score Interpretation and Use*

Grade inflation is present when grades increase over time without a coinciding increase in educational achievement. Because grades are used alongside ACT scores in admissions, course placement, and retention decisions, changes in grading practices affect how each measure contributes to the validity arguments for ACT score use. This section summarizes the current evidence on grade inflation among ACT-tested students and its implications for ACT score interpretation and use across the domains covered in Sections 7.2, 7.3, and 7.4.

#### Historical Context

Early work by Ziomek and Svec (1995) and Woodruff and Ziomek (2004a) documented modest but significant grade inflation among ACT-tested students from 1990 through 2003; HSGPA increased over this period without a corresponding increase in ACT Composite scores. A subsequent study by Zhang and Sanchez (2013) examining data from 2004 to 2011 found no overall pattern of inflation across that 8-year period, although grade inflation and deflation were observed at individual high schools.

#### Recent Trend Evidence (2010–2022)

Sanchez and Moore (2022) examined HSGPA change for 4,393,119 ACT-tested graduates across seven cohorts spanning 2010 through 2021, using hierarchical linear models with students nested within schools. The average HSGPA increased from 3.22 in 2010 to 3.39 in 2021 (a 0.17 grade-point increase), while mean ACT Composite scores declined from 21.0 to

20.3 over the same period. The share of students with an A-average rose from 40.3% in 2010 to 54.9% in 2021, while the share of B-average students fell from 46.8% to 36.2% (see Table 7.8). Grade inflation accelerated sharply during the 2020 and 2021 cohorts, coinciding with the onset of the COVID-19 pandemic, and was greatest at low and moderate ACT score levels. Grade inflation was also greater at schools serving higher percentages of students eligible for free and reduced-price lunch and higher percentages of students of color.

**Table 7.8.** Mean HSGPA and Mean ACT Composite Score, ACT-Tested Public School Graduates, 2010–2021

| Cohort year | Mean HSGPA | Mean ACT Composite | % with A average | % with B average | % with C average |
|---|---|---|---|---|---|
| 2010 | 3.22 | 21.0 | 40.3% | 46.8% | 12.0% |
| 2012 | 3.24 | 21.1 | 41.8% | 46.5% | 11.0% |
| 2014 | 3.25 | 21.0 | 42.3% | 45.6% | 11.2% |
| 2016 | 3.22 | 20.8 | 41.8% | 45.0% | 12.0% |
| 2018 | 3.28 | 20.8 | 42.2% | 46.0% | 10.6% |
| 2020 | 3.37 | 20.6 | 52.6% | 38.1% | 8.4% |
| 2021 | 3.39 | 20.3 | 54.9% | 36.2% | 8.1% |

*Note.* Percentage columns may not sum to 100% because a small share of students have D/F averages. ACT Composite means are the published graduating-class profile means (all students who tested in the given year).

A companion study, Sanchez (2023), extended the analysis to the subject-area level using data from 6,871,894 ACT-tested public high school graduates in the 2010–2022 cohorts. Subject-area GPAs rose between 0.11 and 0.17 grade points over this period, while corresponding ACT subject score means declined between 0.2 and 1.2 scale score points. After accounting for student and school characteristics using hierarchical linear modeling, the study found that inflation was present in all four core subject areas but was largest for math, followed by science, English, and social studies (see Table 7.9). Inflation also varied by student demographic group and school racial and economic composition; Black students experienced the greatest grade inflation across all four subjects.

**Table 7.9.** Change in Mean Subject-Area GPA and Mean ACT Subject Score Between 2010 and 2022, ACT-Tested Public School Graduates

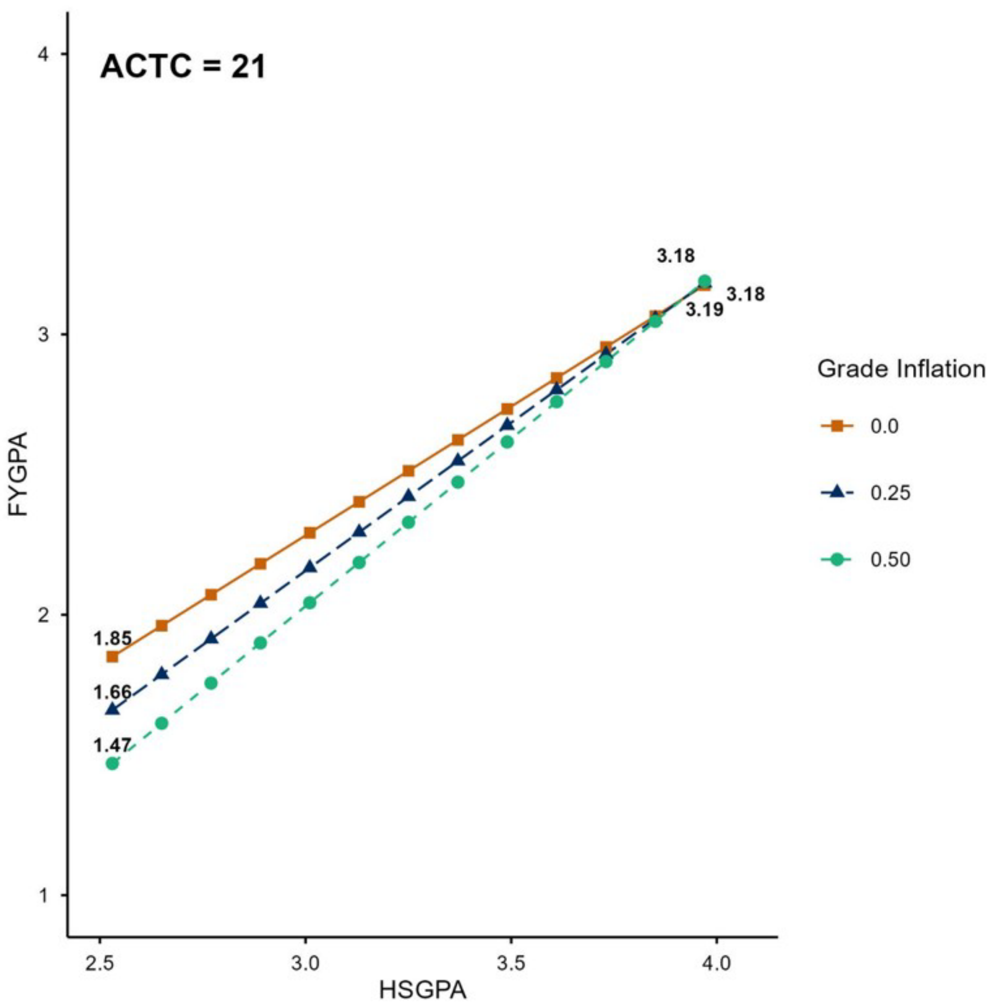
| Subject | 2010 mean GPA | 2022 mean GPA | Δ Observed GPA | Δ Adjusted GPA (HLM) | 2010 ACT | 2022 ACT | Δ ACT |
|---|---|---|---|---|---|---|---|
| Math | 3.19 | 3.36 | +0.17 | +0.30 | 21.4 | 20.2 | −1.2 |
| Science | 3.28 | 3.45 | +0.17 | +0.24 | 21.3 | 20.9 | −0.4 |
| English | 3.30 | 3.41 | +0.11 | +0.22 | 20.9 | 20.2 | −0.7 |
| Social studies | 3.39 | 3.52 | +0.13 | +0.18 | 21.7 | 21.5 | −0.2 |

*Note.* Adjusted GPA change is from a hierarchical linear model controlling for student demographics, course-taking, and school composition. ACT subject score for social studies refers to ACT reading, consistent with the source report.

### Implications for the Predictive Validity of HSGPA

Recent evidence suggests that grade inflation does more than complicate the interpretation of HSGPA; it also measurably affects the predictive validity of HSGPA for college outcomes. Sanchez (2025a) used data from 8,617 students in the 2022 enrolled class at 10 public 4-year institutions in one state with a statewide ACT contract. The study estimated a school-level grade inflation index from cohort-over-cohort HSGPA changes and then used this index as a predictor in a hierarchical linear model of first-year college GPA (FYGPA). Students from higher-inflation schools had systematically lower FYGPA than their HSGPA would have predicted (see Figure 7.3), and the addition of ACT Composite score to the prediction model substantially reduced this overprediction. Although based on data from a single state, these findings are consistent with the role of standardized test scores in mitigating the validity threat introduced by variable grading standards.

**Figure 7.3.** Predicted FYGPA by HSGPA at Three Levels of School Grade Inflation, ACT Composite = 21



### Implications Across Sections 7.2–7.4

This body of evidence on grade inflation has implications for each of the score uses discussed in the remainder of this chapter. For admissions decisions (Section 7.2), rising HSGPA paired with declining ACT Composite scores weakens the shared signal between the two measures; Chen and Sanchez (2024) documented a related shift in the relative predictive strength of HSGPA and ACT Composite score for postsecondary enrollment across the 2010–2021 cohorts. For course placement (Section 7.3), Sanchez (2024a) found that the relationship between HSGPA and placement in developmental courses weakened across cohorts, whereas the ACT Composite score showed a more stable relationship with developmental placement. For retention and degree completion (Section 7.4), Sanchez (2024b) reported declining HSGPA-based predictions of first-year credit hours earned across the 2018–2022 cohorts, while ACT-based predictions remained relatively stable. Collectively, these findings support the continued use of multiple measures, particularly the ACT Composite score alongside HSGPA, when making high-stakes student decisions.

## *7.1.13 Validity Evidence Supporting the ACT Enhancements*

In April 2025, ACT implemented enhancements to the ACT test by reducing testing time, making science optional, and updating the content blueprints while preserving the meaning of the scale and the College Readiness Benchmarks. Allen et al. (2026) documents the design rationale and the validity evidence plan for the enhanced ACT. In updating the test, ACT shortened the English test from 75 items to 40 items plus 10 embedded field-test items, the mathematics test from 60 items to 41 items plus 4 field-test items, the reading test from 40 items to 27 items plus 9 field-test items, and the science test from 40 items to 34 items plus 6 field-test items. The total number of scored items dropped from 215 to 142, and the total testing time for the ACT including science dropped from 175 to 165 minutes. If science is omitted, students complete a 125-minute test with 108 scored items. Beginning in April 2025, ACT began calculating the ACT Composite score based on English, mathematics, and reading (EMR); this replaced the legacy four-section Composite, which included science. The ACT STEM and ACT ELA combined scores remained the same, as did the 1–36 scale, the College Readiness Benchmarks, and the optional writing test.

The design of the enhanced ACT borrows from the assessment design science framework as well as a dual validity loop that pairs content and predictive validity to create test blueprints wherein content, response process, and internal structure evidence are gathered before operational launch, with further data collection thereafter. Initial validity evidence supporting interpretations of enhanced ACT scores includes findings from the June 2024 linking study (Li et al., 2025) and the October 2024 mode comparability study (Allen et al., 2026, Chapter 4), along with reviews conducted by external subject matter experts and multiple cognitive laboratory studies. The results demonstrate that the enhanced ACT measures the same underlying constructs as the legacy ACT while maintaining continuity of the score scales and supporting comparable score interpretations. Confirmatory factor analysis provided evidence of structural equivalence across forms and sections (RMSEA 0.02 to 0.03, CFI 0.95 to 0.99). The reliability coefficients for the English, mathematics, reading, science, EMR Composite, and EMRS Composite were .88, .88, .84, .85, .94, and .96, respectively. Slightly larger standard errors of measurement were seen, as expected from a reduction in scored items. For equating, ACT used IRT true-score equating with the Stocking–Lord (1983) scale transformation. Population invariance, evaluated by WRMSD across subgroups, indicated that with the exception of three statistics, all fell within the range of observed values across the 11 legacy ACT equating baselines. Small but statistically significant score differences were observed via mode comparability analysis for English, mathematics, and reading. Post-adjustment means were nearly identical across modes after IRT true-score equating.

Allen and Cruce (2025) supported the interpretations of the enhanced ACT scores using data from the June 2024 linking study ($N$ = 4,584 enhanced ACT, $N$ = 2,298 legacy ACT) and historical predictive samples from the legacy ACT ($N$ = 1,111,776 students at 527 colleges for first-year college GPA; $N$ = 116,918 students at 2,578 colleges for degree completion). Predictive validity for the enhanced ACT was simulated using a reduced item set from legacy administrations as well as rescaling. This simulated ACT Composite correlated .40 with first-year college GPA versus .41 for the legacy composite, .43 with bachelor's degree completion in 6 years versus .44, and .35 with any degree completion within 6 years versus .35. Concurrent validity correlations with HSGPA were .49 for the enhanced EMRS Composite and .48 for the enhanced EMR Composite compared with .50 for the legacy Composite. Differential validity analysis was conducted across gender, race/ethnicity, grade, and test preparation indicators and identified no significant differences between enhanced and legacy validity coefficients. The EMR Composite and the EMRS Composite correlated .99 across 6 independent samples, and the equal percentile concordance was 1:1 for the score range of 14 to 36. Allen and Cruce concluded that enhanced ACT scores carry the same interpretive meaning as legacy ACT scores for the vast majority of uses identified. These results should be interpreted as provisional until postsecondary outcomes data are available from ACT-tested cohorts.

## 7.2 Using ACT Scores to Support College Admissions Decisions

Attending college requires a significant investment of time, money, and other resources by students and parents, as well as by institutions, which want to admit students who will be academically successful. These parties therefore share a common interest in the investment's success, and the decisions they make about college application and admission are crucial to that success.

Academic achievement is one important aspect of success in college, and academic preparation is one critical determinant of academic achievement. In any postsecondary academic curriculum, a certain minimum level of academic skill is required for success; beyond that, better academic preparation usually results in greater academic success. Therefore, it is appropriate to take students' academic preparation into account when making admissions decisions.

Academic success during a student's college career requires at least a minimal level of academic success in the first year. Some students experience significant academic difficulties in their first year but have satisfactory levels of achievement in subsequent years. Students whose academic difficulties in their first year cause them to leave college cannot be considered academically successful overall. Thus, the likelihood of academic success in the first year is a reasonable factor to consider when making admissions decisions. Because the ACT measures mastery of high school course content, which includes the academic skills needed to succeed in typical first-year college courses, it is appropriate for use in admission.

Although the ACT measures important academic skills needed for success in college, no test can measure all relevant academic skills. Therefore, it is advisable to supplement ACT scores with other academic information, such as courses taken and grades earned in high school, when making admissions decisions.

Moreover, academic preparation is only one determinant of academic success in college, albeit an important one. Nonacademic characteristics such as motivation, interests, and goals can also influence academic success. Therefore, admissions decisions should take into account students' noncognitive characteristics as well as their academic skills. On the ACT, the Student Profile section and the Interest Inventory both provide information on students' background characteristics, goals, and vocational interests. As described above, measures of social-emotional learning competencies may also provide information about students' noncognitive characteristics associated with academic success in college.

Finally, there are other outcomes of postsecondary education (e.g., students' appreciation of culture, their intellectual curiosity, their ability to work with people holding differing opinions) that are not strictly academic in nature but may be considered important educational outcomes. If an institution can define and defend its nonacademic goals and collect information on student characteristics related to them, then such information could also be used in making admissions decisions. Of course, nonacademic characteristics must be validated as predictors of the achievement of nonacademic goals, just as test scores must be validated as predictors of the achievement of academic goals.

### *7.2.1 Statistical Relationships Between ACT Scores and First-Year College GPAs*

If the ACT test measures characteristics important to success in the first year of college, and if first-year grades are reliable and valid measures of undergraduate academic performance, then there should be a statistical relationship between ACT scores and first-year grades. Therefore, a crucial aspect of any validity argument for using ACT scores in making admissions decisions is the strength of the statistical relationships between the test scores and first-year grades.

#### Correlations as Validity Evidence

The Pearson correlation coefficient measures the strength of the linear relationship between two variables, such as college GPA and a test score. The absolute value of the correlation coefficient ranges between 0 and 1, with 0 indicating no relationship and 1 indicating a perfect linear relationship. A correlation near 0 would indicate that the relationship between test scores and GPA is too weak for the test to be used for college admission.

### Measurement Error and Restriction of Range

Two factors attenuate the observed correlation between ACT scores and GPA: measurement error and range restriction. Measurement error effectively places a cap on the observed correlation between two measures because the correlation cannot exceed the square root of the product of the reliabilities of the two measures. Corrections for measurement error in test scores are not made when determining the operational validity of a test since the observed test scores are used in practice. However, corrections for measurement error in course grades or GPA permit an estimation of what the validity of a predictor variable would be if the criterion were measured perfectly. As an example, if the observed correlation between the ACT Composite score and FYGPA were .38, the reliability estimate for FYGPA were .81, and the reliability of ACT scores were 1.0 (no correction for unreliability), the validity coefficient for ACT Composite scores would increase from .38 to .42 ($.38/\sqrt{.81 \times 1} = .42$).

Range restriction in variables also reduces the correlation between predictor and criterion measures, and it is a concern in most institutional validity studies. Specifically, a correlation between test scores and college grades estimated from enrolled students whose academic skills were considered in admitting them will understate the theoretical correlation in the entire applicant population. This statistical problem exists at all postsecondary institutions whose admissions decisions take into account applicants' academic skills. On the other hand, if a college did not use test scores or other measures of applicants' academic skills in making admissions decisions, then applicants with low test scores could enroll as well as those with high test scores. In this situation, the correlation between the students' test scores and their grades would most likely be higher than if the college used test scores in making admissions decisions (Whitney, 1989).

With data from 50 postsecondary institutions, a validity study demonstrated the effects of range restriction on correlations between ACT scores and FYGPA and between HSGPA and FYGPA (Westrick et al., 2015). The correction for range restriction increased the estimated mean correlation between ACT Composite score and FYGPA from .38 to .51, and it increased the estimated mean correlation between HSGPA and FYGPA from .47 to .58 (Table 7.10). Note that the validity coefficients for ACT Composite score and HSGPA were somewhat variable across institutions—as indicated by the 90% credibility intervals—yet the corrected correlation between socioeconomic status (SES) and FYGPA did not vary across institutions.

**Table 7.10.** Meta-Analysis of Multi-Institutional Data—Correlations With FYGPA, Overall Analyses

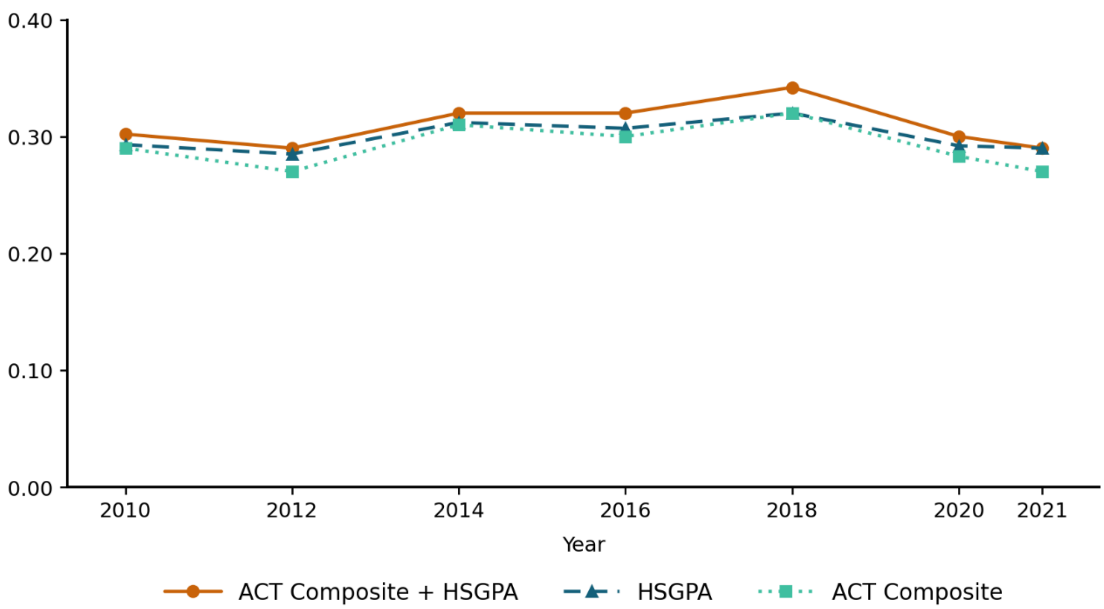
| Predictors | *k* | *N* | Mean observed *r* | $SD_r$ | Estimated mean $\rho$ | $SD_\rho$ | 95% CI | 95% CrI |
|---|---|---|---|---|---|---|---|---|
| ACT Composite | 50 | 169,818 | .38 | .07 | .51 | .05 | .50, .53 | .43, .60 |
| HSGPA | 50 | 150,305 | .47 | .05 | .58 | .06 | .57, .60 | .49, .68 |
| SES | 50 | 139,354 | .12 | .04 | .24 | .00 | .24, .25 | .24, .24 |

*Note.* *k* = number of institutional studies; $SD_r$ = standard deviation of observed correlations; $SD_\rho$ = standard deviation of correlations corrected for artifacts; CI = confidence interval; CrI = credibility interval.

A follow-up study demonstrated that ACT Composite scores provided incremental improvement to the prediction of FYGPA after HSGPA was accounted for (Westrick et al., 2021). Such studies support the use of ACT scores in combination with other measures to help inform college admissions decisions. In that study, after the researchers corrected for range restriction and measurement error in both the predictors and the criterion, the estimated mean multiple correlation for the FYGPA prediction model that included ACT Composite score and HSGPA was .662, compared to .627 for the model that included only HSGPA. The incremental validity finding has been corroborated by other studies, including a study by Allen et al. (2019) that reported an increase in the multiple correlation (from .48 to .52) when ACT Composite score was added to the FYGPA prediction model that included HSGPA.

Recent research has examined whether the predictive validity of college readiness measures has shifted over time. Chen and Sanchez (2024) examined the predictive validity of HSGPA and ACT Composite score for postsecondary enrollment across the 2010–2021 ACT-tested graduating cohorts, matched to National Student Clearinghouse records. Using logistic regression with student and school controls, the study found that the predictive validity of HSGPA for postsecondary enrollment increased steadily from 2010 to 2021, while the predictive validity of ACT Composite score declined over the same period. Prior to 2014, ACT Composite score was the stronger predictor of enrollment; after 2014, HSGPA became the stronger predictor. However, across all cohorts examined, models that included both measures outperformed models using either measure alone (see Figure 7.4). These findings are consistent with the broader shift toward test-optional admissions policies and reinforce the value of multiple measures for understanding college readiness.

**Figure 7.4.** Conditional $R^2$ of HSGPA and ACT Composite Score From 2010 to 2021



Complementing the Chen and Sanchez (2024) study, Sanchez (2024c) used dominance analysis to evaluate the relative importance of HSGPA, ACT Composite score, and demographic factors for predicting FYGPA. Dominance analysis is a regression-based technique that compares the contribution of multiple predictors to $R^2$ across all possible subset regression models, providing a more complete picture of relative predictor importance than standard regression. Using a sample of 7,924 ACT-tested 2021 graduates from 10 public 4-year institutions in one statewide-testing state, the study found that academic measures dominated demographic factors in predicting FYGPA. Family income and gender were the weakest predictors among all predictors examined. Consistent with prior research, ACT Composite score and HSGPA both contributed meaningful unique variance to FYGPA prediction; however, no clear dominance of one academic measure over the other was determined (see Table 7.11). Although limited to a single state, these findings suggest that conclusions about the relative importance of HSGPA and ACT Composite score depend in part on the modeling approach used.

**Table 7.11.** General Dominance Statistics for Predictors of FYGPA

| Predictor | $R^2$ contribution | Rank |
|---|---|---|
| HSGPA | 0.11 | 1 (tied) |
| ACT Composite score | 0.11 | 1 (tied) |
| Race/ethnicity | 0.03 | 3 |
| Family income | 0.02 | 4 |
| Gender | 0.01 | 5 |

*Note.* Adapted from Sanchez (2024c). $R^2$ contribution values are averaged across 20 multiply imputed datasets and represent each predictor's contribution to modeling $R^2$ across all possible subset regression models. HSGPA and ACT Composite score were tied as the most important predictors; no clear dominance of one academic measure over the other was determined.

### *Multiple Scores and Superscoring*

Another factor that can affect validity coefficients is the choice of which ACT scores to include in an analysis for students who take the ACT more than once. Postsecondary institutions report using a variety of composite scoring methods for applicants with multiple test records, including averaging and taking the maximum section test scores across test occasions ("superscoring"). A recent study evaluated the validity of using ACT scores obtained from various scoring methods across test administrations (average, highest, most recent, and superscoring) to identify students who are likely to succeed in their first year of college (Mattern et al., 2018). The study found that, compared to the other options, superscores were as predictive—if not more predictive—of first-year grade point average and resulted in the least amount of differential prediction when the researchers statistically controlled for the number of times tested. Specifically, the differential prediction results suggested that first-year grades for examinees who tested more often tended to be underpredicted by ACT scores. That is, retesters performed better in college than their scores suggested they would. However, this prediction error was lower when superscoring was used instead of the other scoring methods.

A follow-up study examined the relationship between different scoring methods and another important measure of college success: degree completion (Radunzel & Mattern, 2020). Consistent with the earlier study, this study indicated that superscores were as predictive of completing a college degree in a timely manner as the other scoring methods. The study also revealed that the likelihood of completing a degree for students who tested more often was underpredicted but that superscoring resulted in the least prediction error when prediction accuracy was examined based on the number of times students tested. These findings held for both the 2- and 4-year institution samples. These two studies provide evidence supporting ACT's new practice of reporting the ACT Superscore, since they suggest that selecting students' best scores from any test attempt results in the most predictive indicator of a student's future academic success.

### Decision-Based Statistics as Validity Evidence

The correlation coefficient is used more often than any other statistic to summarize the results of predictive validity studies. As an index of the strength of the linear relationship between first-year college grades or GPAs and admissions or placement measures, a correlation coefficient can lend credibility to a validity argument. However, it does not directly measure the degree to which admissions or placement measures correctly identify students who are academically prepared for college coursework. Rather, the correlation coefficient reflects the accuracy of prediction across all values of the predictor variables. Of greater interest to educators who must evaluate admissions or placement systems is the correctness of the decisions made about individual students and their estimated chances of success. In this section, logistic regression and decision-based statistics are described as alternative methods for summarizing the results of predictive validity studies. Studies presented in this section demonstrate the use of this method for making admissions and course placement decisions.

Suppose success in the first year of college can be defined in terms of some measurement that is obtainable for each student; for example, success might be defined as a student completing the first year with a GPA of C or higher in a common subset of first-year courses. There are therefore four possible outcomes of the admissions decision for a particular student:

A. True positive: The student is permitted to enroll in the college and is successful there. (Correct decision)

B. False positive: The student is permitted to enroll in the college and is not successful there. (Incorrect decision)

C. True negative: The student is not permitted to enroll in the college and would not have been successful there. (Correct decision)

D. False negative: The student is not permitted to enroll in the college but would have been successful there. (Incorrect decision)

The sum of the proportions associated with outcomes A and C is the proportion of correct admissions decisions.

Note that outcomes A and B can be directly observed in existing admissions systems, but outcomes C and D cannot. In principle, the proportions associated with all four outcomes could be estimated by collecting admissions measures (e.g., admissions test scores) for every student while permitting everyone to enroll in the college regardless of test score. Some of these students would be successful in the college, and others would not; the relationship between the probability of success and the admissions measures could then be modeled using statistical methods. From the estimated conditional probabilities of success for given values of the admissions measures, estimates of the probabilities of Outcomes A–D could be calculated.

In most institutions, of course, this kind of experimentation is not done because students with low probabilities of success are generally not admitted or do not select the college. Therefore, first-year outcomes are not available for these students, and the relationship between their probability of success and their admissions measures must be estimated by extrapolating relationships estimated from the data of students who actually enrolled in the college. The assumption being made is that the conditional probability of success given the selection variable(s) is the same for the nonenrolled applicants as for the enrolled students. This assumption is analogous to that used for the traditional adjustment of correlations for restriction of range, which requires that the applicant and enrolled student groups have the same conditional mean and variance functions (e.g., Lord & Novick, 1968). Research at ACT has shown that accurate extrapolations can usually be made from moderately truncated data (Houston, 1993; Schiel & Harmston, 2000; Schiel & Noble, 1992).

It is possible to relate a correlation coefficient to the function for the conditional probability of success, but a number of strong statistical assumptions are required. A more straightforward way to estimate the probability of success is to dispense with correlation coefficients altogether and model it directly. For example, one could use the logistic regression model

$$\hat{P}[W = 1|X = x] = \frac{1}{1+e^{-\hat{a}-\hat{b}x}}, \quad (1)$$

where $W$ = 1 if a student is successful in college,

$W$ = 0 if a student is not successful in college, and

$X$ is the student's admissions test score.

An example of an estimated logistic function is the curve labeled "Probability of C or higher" in Figure 7.5. Note that the probability of a C or higher ranges from .05 to .99 depending on the test score. Note also that this curve is S-shaped and that its maximum slope occurs at the test score of 20. In logistic regression, the point at which the maximum slope occurs is called the inflection point, and the slope of the curve at this point is proportional to the coefficient $\hat{b}$ in Equation 1. Therefore, a larger $\hat{b}$ value corresponds to a steeper slope, which indicates that the test is better at discriminating between students who will and will not succeed (for students with test scores near the inflection point).

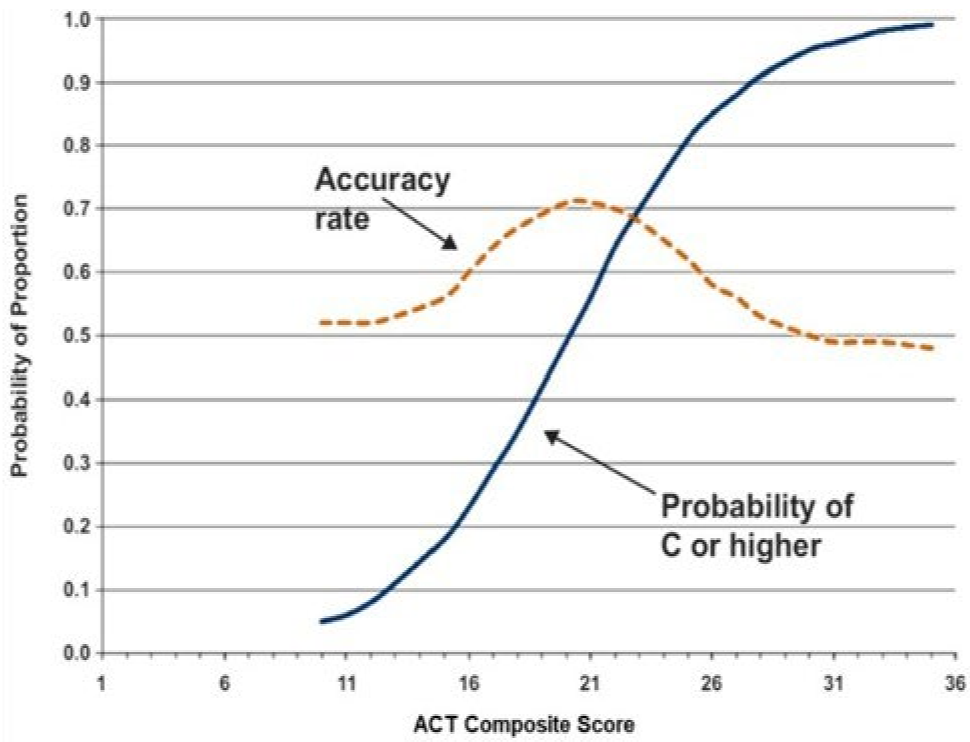
The estimated $\hat{a}$ and $\hat{b}$ coefficients in Equation 1 can be calculated by iterative least squares procedures. Given the previous discussion, the coefficient $\hat{b}$ should be positive and significantly different from zero. A coefficient near zero would result in a flat curve for the conditional probability of success.

Once estimates $\hat{a}$ and $\hat{b}$ have been obtained, estimated probabilities for the four outcomes can be calculated easily. For example, if 16 is the cutoff score on *X* for being admitted to an institution, then the probability of a true positive (Outcome A) can be estimated by

$$\hat{P}[A] = \frac{\sum_{x \geq 16} \hat{P}[W=1|X=x]n(x)}{N}, \tag{2}$$

where $\hat{P}[W=1|X=x]$ is Equation 1 calculated from the estimates $\hat{a}$ and $\hat{b}$, $n(x)$ is the number of students whose test score is equal to $x$, and $N$ is the total number of students in the sample. At institutions with existing admissions systems, the conditional probabilities $\hat{P}[W=1|X=x]$ in Equation 1 are calculated from data for students who enrolled in the institution. The probability $\hat{P}[A]$ in Equation 2, however, is calculated from the test scores of all students, both those who were admitted and those who were not. The probabilities for Outcomes B, C, and D can be estimated in a similar way.

**Figure 7.5.** Probability of C or Higher FYGPA and Accuracy Rate




It should be noted that admissions decisions are usually made based on several measures. For the purpose of illustrating how the accuracy of admissions decisions can be estimated, the example uses a simplified model based on a cutoff score on a single admissions test. Students scoring at or above the cutoff score would be admitted; students scoring below the cutoff score would not. ACT does not advocate making admissions decisions based on a single measure; this example is for illustration only. Results provided later in this section illustrate how the logistic regression model may be generalized to multiple measures.

Once the estimates $\hat{P}[A]$ and $\hat{P}[C]$ are obtained, the percentage of correct admissions decisions (the accuracy rate) is estimated as $\hat{P}[A]$ + $\hat{P}[C]$ multiplied by 100. An illustration of estimated accuracy rates for different test scores is given in Figure 7.5 as a proportion. Note that the maximum accuracy rate (0.71) occurs at the inflection point in the graph of the probability of success (i.e., near a score of 20). This score is referred to as the optimal cutoff score, the score that maximizes the percentage of correct admissions decisions.

The accuracy rate value corresponding to the lowest obtained test score represents the overall percentage of students who would be predicted to succeed in college if the test were not used for admissions decisions. The difference between the maximum accuracy rate and the accuracy rate for the lowest test score (the increase in accuracy rate) is an indicator of the effectiveness of the test for making admissions decisions. This statistic shows the increment in the percentage of correct admissions decisions due to the use of the test. Large increases in accuracy rate correspond to a greater contribution by the test to increasing the rate of correct admissions decisions. Note that a selection variable has incremental accuracy if and only if its probability-of-success curve crosses 0.5 somewhere.

The ratio of true positives, $\hat{P}[A]$, to the sum of true positives and false positives, $\hat{P}[A]$ + $\hat{P}[B]$, multiplied by 100 indicates the estimated percentage of students who would be successful out of those who would be admitted based on particular admissions criteria. This ratio is called the success rate. Like the probability of success, the success rate should increase as scores on the admissions measure increase. The incremental success rate associated with a selection variable is the difference between the success rate associated with admitting applicants at or above the specific cutoff score and the success rate associated with admitting all applicants (i.e., the base success rate for the lowest test score).

### College Admission Validity Evidence Using Decision-Based Statistics

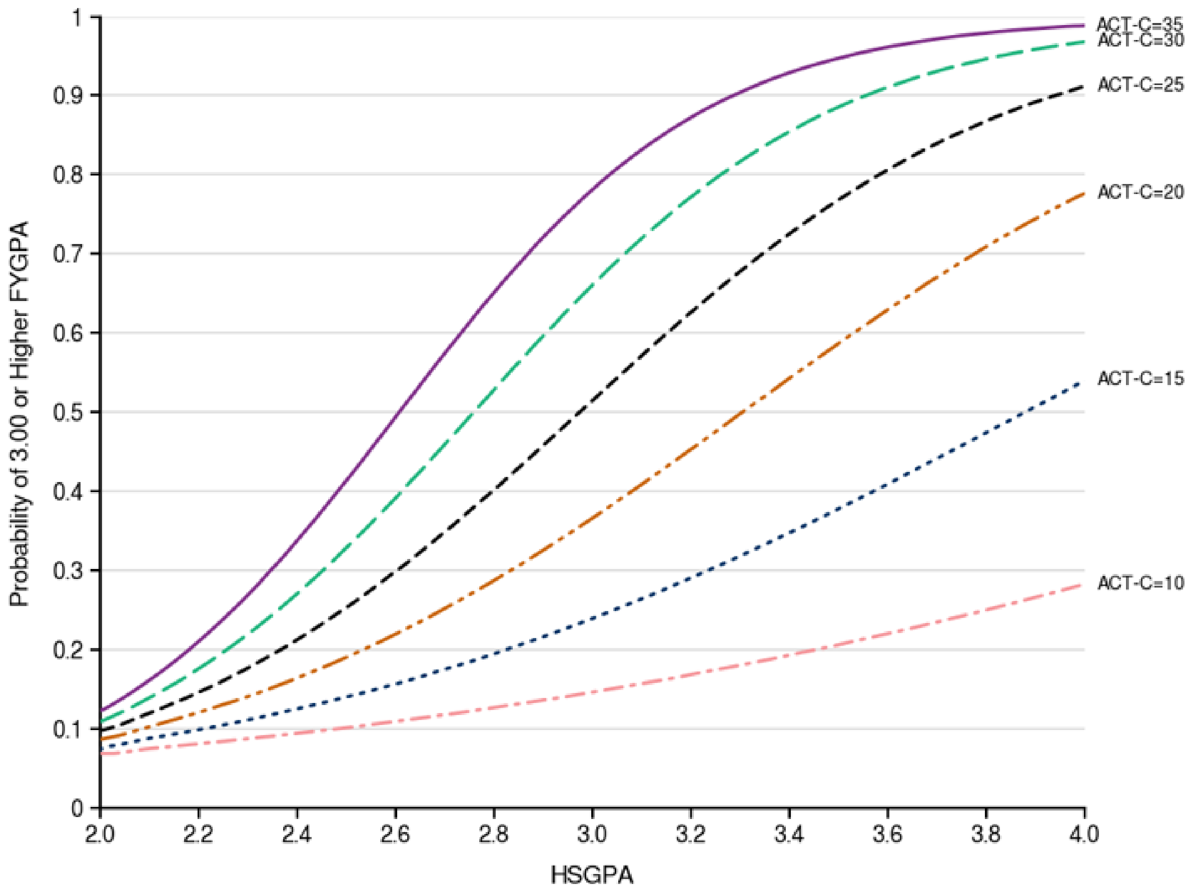
A majority of postsecondary institutions use standardized test scores in combination with high school grades or rank for making admissions and course placement decisions (Clinedinst, 2015). This practice is supported by research demonstrating the validity of using multiple measures for making college admissions and placement decisions (e.g., Noble et al., 1995; Noble & Sawyer, 2002) and the reality that no test can measure all the skills and knowledge needed for success in college. Using multiple measures can increase content coverage and, as a consequence, increase the accuracy of admissions decisions.

The usefulness of a selection variable for admission to college depends in large part on its predictive power, but it also depends on admissions officers' goals, which are aligned to their institutions' larger goals of educating students successfully. Usefulness also depends on other factors, such as applicant self-selection and institution selectivity. To gauge the usefulness of a selection variable, one must specify the goals of using that variable. Two common goals related to academic achievement are (a) maximizing academic success among enrolled students and (b) accurately identifying applicants who would be academically successful at the institution and enrolling as many of them as possible.

These goals may seem similar, but they are not identical. The first is related to the proportion of applicants who would succeed academically if they enrolled (i.e., the success rate). The second is related to the proportion of applicants that an institution correctly identifies as likely to succeed or fail (i.e., the accuracy rate). Both goals, however, pertain only to institutions with some degree of selectivity in their admissions policies rather than to institutions with open-admission policies.

A study was conducted to evaluate the usefulness of ACT Composite score and HSGPA for college admissions decisions using the decision-based statistics discussed in the previous section (Sawyer, 2010). Using data from 192 four-year postsecondary institutions, the study evaluated whether using ACT Composite score for selection increased the success rate and accuracy rate beyond what would result if the institution did not use ACT Composite score. For example, Figure 7.6 shows the probability of earning a FYGPA of 3.0 or higher given different values of HSGPA and ACT Composite score. The model coefficients for the main effects (HSGPA and ACT Composite score) and their interaction were statistically significant ($p < .001$). The interaction indicated that the slope of the probability-of-success curve increased as ACT Composite score increased. One interpretation of the interaction term is that HSGPA is more predictive for students with higher ACT Composite scores than for students with lower ACT Composite scores. This figure also illustrates that the probability of earning a FYGPA of 3.0 or higher varies dramatically among students with the same HSGPA but different ACT Composite scores. Among students with a 4.0 HSGPA, students with an ACT Composite score of 15 have a probability of .54, whereas students with an ACT Composite score of 30 have a probability of over .95. Even in less extreme cases, the results illustrate that ACT Composite scores meaningfully discriminate among students with the same HSGPA in terms of success in the first year of postsecondary studies.

**Figure 7.6.** Probabilities of Earning a 3.0 or Higher FYGPA and Being Retained Through the First Year Based on HSGPA and ACT Composite Score




Results from the Sawyer (2010) study were consistent with those from an earlier study by Noble and Sawyer (2002). Results from both studies suggest that HSGPA by itself is better than ACT Composite score by itself for some, but not for all, degrees of selectivity and definitions of success. The ACT Composite score is more useful than HSGPA in some situations—for example, when an institution is interested in high levels of success. In most scenarios, using both high school grades and test scores jointly is better than using either by itself. When using both variables, institutions should take into account the HSGPA by ACT Composite score interaction effect as well as the main effects.

Postsecondary institutions seek high achievement for their students and want to admit students who have a good chance of succeeding in college. The results from these studies suggest that ACT Composite scores provide differentiation across levels of achievement in terms of students' probable success during their first year in college.

### *7.2.2 Differential Prediction in First-Year College GPA Among Student Groups*

Differential prediction occurs when students who have the same test scores but belong to different population groups have different probabilities of success. One of the effects of differential prediction is that, if an institution makes decisions using cutoff scores based on students' probability of success, different observed success rates can result for different population groups. For example, predictive correlations could differ among the groups. Another possibility could be that the proportion of admitted applicants who are successful (success rate) and the proportion of correct admissions decisions (accuracy rate) could differ. Any such differences may result from differential validity.

### Differential Prediction by Race/Ethnicity, Gender, and Family Income

A study by Sanchez (2013) investigated the differential effects on student groups of using ACT Composite score and HSGPA to make admissions decisions. Student characteristics included race/ethnicity, gender, and income. For each student group, Sanchez examined the effect of using a total group cut score for ACT Composite score, HSGPA, or both on predicting first-year college grade point average (FYGPA).

The results indicated that the probability of earning a 2.5 or 3.0 FYGPA increased as ACT Composite score or HSGPA increased, and this was true for White, African American, and Hispanic students (Figures 7.7 and 7.8). For the two FYGPA levels (2.5 and 3.0), White students had higher estimated probabilities of success than African American and Hispanic students over most of the ACT Composite score and HSGPA scales, and Hispanic students tended to have higher estimated chances of success than African American students. When differential prediction was apparent, it tended to be of greater magnitude when HSGPA was used as the academic predictor than when ACT Composite score was used (Figure 7.8 versus Figure 7.7). This was particularly notable for African American students scoring above a HSGPA of about 3.0. This suggested that a total-group HSGPA model considerably overestimates the chances of success for African American and Hispanic students with a high HSGPA.

**Figure 7.7.** Estimated Probabilities of Achieving Specific FYGPA Levels Based on ACT Composite Score, by Race/Ethnicity

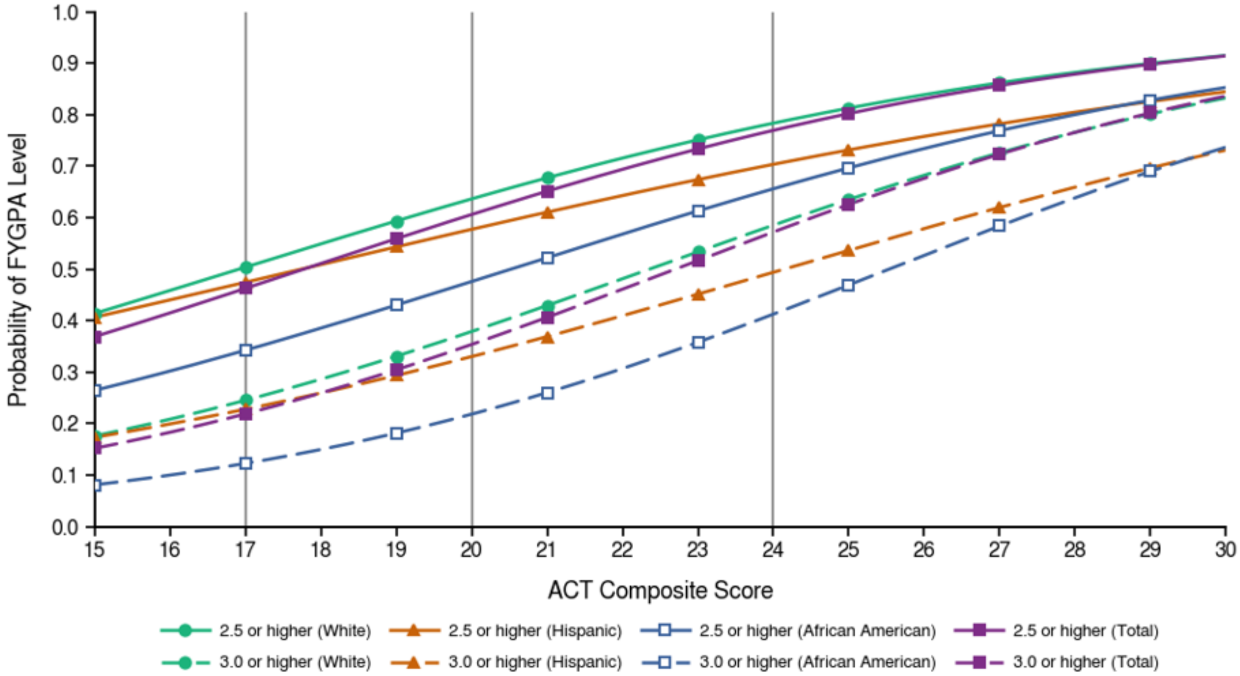
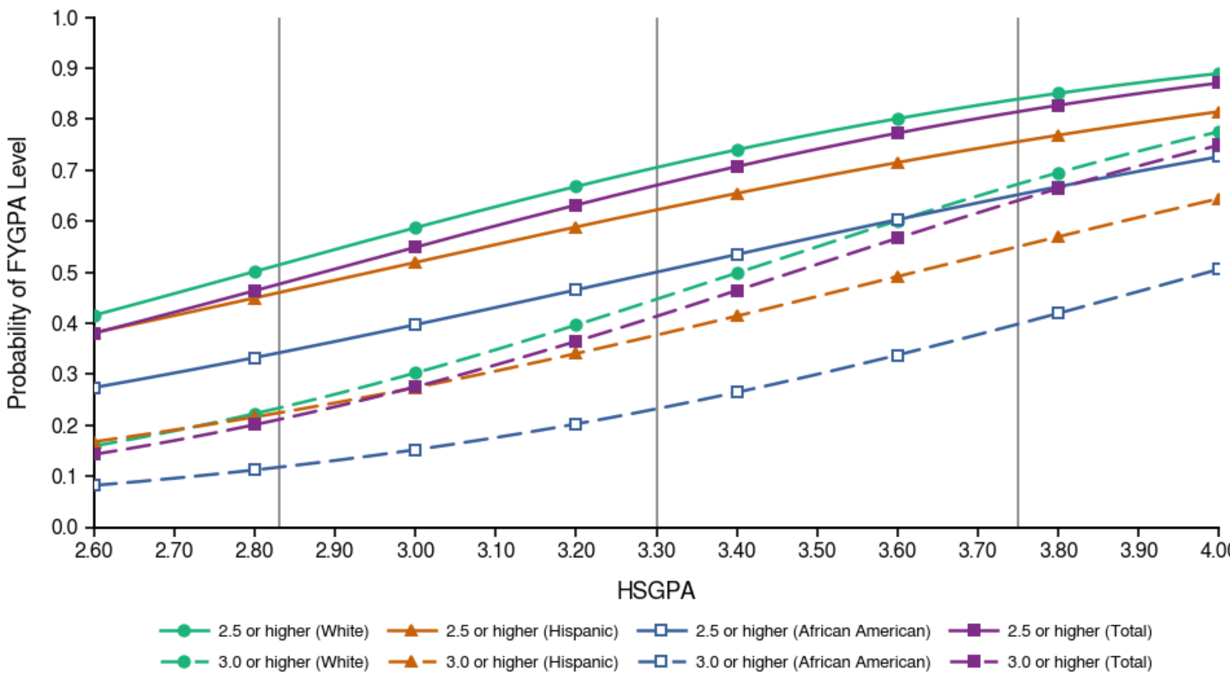



**Figure 7.8.** Estimated Probabilities of Achieving Specific FYGPA Levels Based on HSGPA, by Race/Ethnicity



Consistent with Figures 7.7 and 7.8, the median probabilities of success across institutions based on a total-group cutoff for racial/ethnic groups tended to show a pattern of underprediction for White students and overprediction for both Hispanic and African American students (Table 7.12). The joint ACT Composite score and HSGPA model tended to produce the most favorable accuracy rates and success rates, on average, across the racial/ethnic groups. For the 3.0 or higher FYGPA success level, median accuracy rates were highest for African American students and lowest for White students. Moreover, the increase in accuracy rates associated with using ACT Composite score and HSGPA jointly as predictors was greater for African American and Hispanic students than for White students.

**Table 7.12.** Median Statistics for Predicting 3.0 FGYPA or Higher by Race/Ethnicity Across 236 Institutions

| Predictor | Total-group cutoff | Race/ ethnicity | Median | | | |
|---|---|---|---|---|---|---|
| | | | Probability of success | Maximum accuracy rate | Accuracy rate increase | Success rate |
| ACT Composite score | 23 | White | .54 | 71 | 25 | 68 |
| | | African American | .36 | 86 | 71 | 46 |
| | | Hispanic | .45 | 78 | 56 | 53 |
| HSGPA | 3.4 | White | .52 | 72 | 22 | 68 |
| | | African American | .27 | 81 | 64 | 37 |
| | | Hispanic | .42 | 75 | 49 | 52 |
| ACT Composite score & HSGPA | NA | White | .51 | 75 | 30 | 70 |
| | | African American | .32 | 87 | 73 | 48 |
| | | Hispanic | .43 | 81 | 61 | 55 |

*Note.* Multiple combinations of ACT Composite score and HSGPA correspond to a .50 probability of success in the joint models.

Overall, the study revealed that across student groups, the joint use of ACT Composite score and HSGPA resulted in greater prediction accuracy than either predictor alone. Furthermore, the use of a total-group cutoff score for both ACT Composite score and HSGPA, individually and in combination, slightly overpredicted the probability of success of Hispanic and African American students, males, and students from low-income households. These findings indicate that using test scores, alone or in combination with HSGPA, does not underpredict the first-year performance of African American, Hispanic, or low-income students; if anything, a total-group model tends to slightly overpredict their performance. Because overprediction may also reflect limitations of the grade criterion and factors the model does not capture, these results should not be read as establishing that the broader admissions process is free of disadvantage for these groups.

These results are further corroborated by findings from a parallel study (Radunzel & Noble, 2013) that examined the differential effects on student demographic groups of using ACT scores and HSGPA for predicting long-term college success through degree completion. In general, research indicates that the use of multiple measures helps capture a more holistic view of student readiness and improves prediction of postsecondary success. As a case in point, results from a study by Mattern et al. (2017) suggested that including noncognitive measures such as academic discipline (the amount of effort a student puts into schoolwork and the degree to which a student sees himself or herself as hardworking and conscientious) in a prediction model that already included ACT Composite score and HSGPA increased the variance accounted for in FYGPA and reduced the amount of differential prediction by gender.

### Differential Prediction for Students Testing With Accommodations

Since the enactment of the Individuals with Disabilities Education Act (IDEA) in 1975, the total percentage of disabled students enrolled in public schools has increased from 8.3% (1976–1977) to 8.8% (2004–2005), and the percentage has hovered around 13% to 14% from 2005 to 2019 (de Brey et al., 2021; Snyder & Dillow, 2013). The number of students who elect to take the ACT under special conditions continues to grow. Accommodations for eligible students with disabilities are discussed in Chapter 4. These include, but are not limited to, large-type editions, braille editions, prerecorded audio and video American Sign Language (available via URL), and reader's script administration.

Table 7.13 shows average ACT scores for a sample of 436,695 students who tested with accommodations and 7,252,520 students who tested without accommodations between September 2016 and July 2020. On average, accommodated students achieved lower ACT scores than non-accommodated students, with accommodated students earning an average ACT Composite score of 18.4 and non-accommodated students earning an average ACT Composite score of 20.5. However, there were large differences in performance by disability category.

Because of the growing number of students with disabilities, it is important to demonstrate that a student's ACT scores and HSGPA are valid predictors of college success not only for students who tested under regular conditions but also for students who tested with accommodations. Several prior studies have demonstrated the validity of ACT Composite score and HSGPA in predicting the FYGPA of students with disabilities who received a testing accommodation (Laing & Farmer, 1984; Ziomek & Andrews, 1996). A more recent study by Huh and Huang (2016) examined this issue.

Huh and Huang (2016) found that ACT test scores obtained during an accommodated administration for students with disabilities are predictive of FYGPA. Moreover, using multiple measures provides a more accurate prediction of special-tested students' chances of succeeding in college. Specifically, this study found that a prediction model that uses both ACT Composite score and HSGPA is a good model for predicting actual college FYGPA for both students testing with accommodations and those testing without accommodations. Full results can be found in ACT Research Report 2016-7.

**Table 7.13.** Average ACT Scores for Accommodated and Non-Accommodated Examinees, by Disability Category

| Disability category | | *n* | Percentage | English | Math | Reading | Science | Composite |
|---|---|---|---|---|---|---|---|---|
| **Neuro-developmental disability** | **Total** | **317,603** | **73%*** | **16.0** | **17.4** | **18.0** | **17.8** | **17.4** |
| | ADHD | 87,626 | 28% | 18.8 | 19.3 | 20.7 | 19.8 | 19.8 |
| | Autism | 18,076 | 6% | 18.5 | 18.6 | 19.8 | 19.5 | 19.3 |
| | Communication disorder | 6,954 | 2% | 14.6 | 16.5 | 16.4 | 16.7 | 16.2 |
| | Intellectual disorder | 9,179 | 3% | 11.2 | 14.3 | 13.1 | 14.0 | 13.3 |
| | Learning disability, math | 16,286 | 5% | 14.3 | 15.0 | 16.3 | 15.3 | 15.4 |
| | Learning disability, reading | 44,479 | 14% | 15.4 | 17.7 | 17.9 | 18.0 | 17.4 |
| | Learning disability, writing | 3,829 | 1% | 15.4 | 17.9 | 17.6 | 18.1 | 17.4 |
| | Learning disability, unspecified | 1,722 | 1% | 13.8 | 15.8 | 15.9 | 15.9 | 15.5 |
| | Motor disability | 431 | 0% | 21.4 | 21.1 | 22.3 | 21.5 | 21.7 |
| | Other neurodevelopmental disability | 71 | 0% | 26.1 | 23.7 | 26.7 | 24.9 | 25.5 |
| | Multiple neurodevelopmental disabilities | 128,950 | 41% | 14.6 | 16.5 | 16.8 | 16.8 | 16.3 |
| **Physical/sensory disability** | **Total** | **32,001** | **7%*** | **19.5** | **19.9** | **21.3** | **20.5** | **20.4** |
| | Hearing impairment | 4,738 | 15% | 15.1 | 17.1 | 17.4 | 17.7 | 16.9 |
| | Motor impairment | 2,014 | 6% | 20.2 | 19.9 | 22.6 | 21.0 | 21.0 |
| | Visual impairment | 4,777 | 15% | 20.1 | 20.4 | 22.5 | 21.0 | 21.1 |
| | Other physical/sensory disability | 19,139 | 60% | 20.3 | 20.5 | 21.7 | 21.0 | 21.0 |
| | Multiple physical/sensory disabilities | 1,333 | 4% | 19.8 | 20.0 | 22.1 | 20.9 | 20.8 |
| **Psychological disability** | **Total** | **28,712** | **7%*** | **20.3** | **19.8** | **22.3** | **20.6** | **20.9** |
| | Anxiety | 10,261 | 36% | 23.3 | 21.9 | 25.4 | 23.1 | 23.6 |
| | Conduct disorder | 9,899 | 34% | 14.9 | 16.0 | 16.7 | 16.3 | 16.1 |
| | Depression | 1,000 | 3% | 21.6 | 20.7 | 23.8 | 21.6 | 22.0 |
| | Other psychological disability | 1,089 | 4% | 21.6 | 21.2 | 23.3 | 21.9 | 22.1 |
| | Multiple psychological disabilities | 6,463 | 23% | 23.4 | 21.8 | 25.5 | 23.0 | 23.5 |
| **Multiple disabilities** | | **54,416** | **12%*** | **20.7** | **20.2** | **22.6** | **21.2** | **21.3** |
| **Other disability** | | **3,963** | **1%*** | **19.1** | **19.6** | **20.8** | **20.1** | **20.0** |

| Disability category | *n* | Percentage | English | Math | Reading | Science | Composite |
|---|---|---|---|---|---|---|---|
| **All accommodated examinees** | 436,695 | — | 17.1 | 18.1 | 19.2 | 18.6 | 18.4 |
| **Non-accommodated** | 7,252,520 | — | 19.7 | 20.2 | 21.0 | 20.5 | 20.5 |

*Percentage of all accommodated examinees. All other percentages are the percentages of examinees within each high-level disability category.

## 7.3 Using ACT Scores to Support Course Placement Decisions

One common use of the ACT is to help place students into first-year college courses. This section summarizes research conducted on the effectiveness of ACT scores for this use.

At many postsecondary institutions, there are two levels of first-year courses: standard courses, in which most students enroll, and remedial or developmental courses for students who are not academically prepared for standard courses. Some institutions may also have advanced or honors courses for exceptionally well-prepared students.

In all cases, one can think of placement as a decision about whether to recommend that a student enroll in an upper-level or a lower-level course. The names upper-level and lower-level may refer variously to standard and developmental courses or to advanced and standard courses. Placement systems typically identify students with a small chance of succeeding in an upper-level course and therefore recommend that they enroll in a lower-level course.

### *7.3.1 Placement Validity Argument Based on ACT Content*

A validity argument for a placement test can be based in part on subject matter content. The full ACT test is intended to measure academic skills and knowledge that are acquired in typical college-preparatory curricula in high school and that are essential for academic success in the first year of college. The content specifications of the ACT are based on the recommendations of nationally representative panels of secondary and postsecondary educators (see Chapter 2). Determining the content alignment between the ACT and a particular course at a given postsecondary institution must, of course, be done by faculty at the institution who know the course content. ACT therefore recommends that faculty and staff review the ACT test content and specifications to determine their relationship to the first-year curriculum as a preliminary step in deciding whether to use the ACT for first-year course placement.

Given that the content of the ACT tests is related to the skills and knowledge required for success in college, and given that course grades are reliable and valid measures of educational performance in the course, there should be a statistical relationship between test scores and course grades. If there is close content alignment between an ACT test and the college course, then it is reasonable to expect that students with higher ACT scores will tend to be more successful in the college course than students with lower ACT test scores. If this expectation of ACT scores is borne out in empirical studies, then it is appropriate to consider using the tests for course placement.

As noted previously, the ACT does not measure all aspects of students' readiness for all first-year college courses. Therefore, it is advisable to consider using measures such as high school coursework and grades, scores on locally developed placement tests, or noncognitive measures in addition to ACT scores when making placement decisions. Feasibility and cost are two key issues in deciding whether and how to use additional measures of academic skills for course placement.

### 7.3.2 Statistical Relationships Between ACT Scores and Course Grades

ACT has collected course grades from postsecondary institutions specifically to examine the effectiveness of the ACT test for placement. Results from these analyses provide validity evidence for using ACT scores for placement.

#### Data and Method

Grade data for several different course types were gathered from entry-level courses at 2-year and 4-year institutions. Within each institution, courses that had at least 50 students who earned a course grade and had an ACT score on the corresponding section test were included in the analyses. The sample for each course was weighted to match the population of ACT-tested enrollees at each institution in terms of gender, race/ethnicity, ACT Composite score level, and HSGPA level.

Logistic regression models were used to estimate probabilities of success for each course for each institution (data permitting). Course success, which was defined as earning a grade of B or higher, was predicted from the relevant ACT score. Only courses with success rates between 20% and 80% and with logistic regression curves that crossed the .50 probability level were retained in the analyses.

At each ACT score, the success and accuracy rates were estimated from the probabilities of success obtained from the logistic regression model (see Section 7.2.1 for descriptions of these statistics). These decision-based statistics were then summarized across institutions by course type. To assess validity, accuracy rates were summarized at the institution-specific optimal cutoff score, which is the ACT cutoff score that, if used for course placement, would provide the most accurate prediction of course success. When examined across a range of possible cutoff scores for a given institution, the accuracy rate typically peaks at a specific score and then decreases as the score increases.

This maximum value, which corresponds to a .50 probability of success, is the optimal cutoff score for a given course. There are four reasons why success was defined as a grade of B or higher rather than C or higher:

1. The statistical model would be unstable if success or failure occurred rarely, and grades below C are fairly uncommon in most courses.

2. If the optimal cutoff score is used for course placement, the least-qualified student allowed into the course has about a 50% chance of success. If success is defined as a grade of C or higher, that means the least-qualified student has about a 50% chance of getting a D or F. It would seem poor policy to place a student into a course if the student had such a high likelihood of needing to repeat the course due to poor grades.

3. The success criterion of B or higher results in grade distributions that more closely follow those currently found in colleges. As noted above, grades below C are fairly uncommon

in most courses. Moreover, the mean FYGPA tends to be closer to 3.0 than to 2.0 in recent studies (Allen & Radunzel, 2016; Radunzel & Noble, 2012; Sawyer, 2013a).

4. Prior studies have shown that students who earn B or higher grades in the first year of college are much more likely to earn a college degree than those who earn lower grades (Allen, 2013).

Validity can also be examined by determining the strength of the relationship between the predictor (ACT scores) and course success. The logistic regression model is defined by intercept and slope coefficients, and the slope indicates the strength of the relationship. To summarize the strength of the relationship, median logistic regression slopes are also provided.

## Results

### *Accuracy Rate*

Table 7.14 provides the summarized results for 17 courses. For all courses, the median accuracy rate at the optimal cutoff score was at least 62% (refer back to Section 7.2.1 for an overview of decision-based statistics). Thus, a typical institution using the ACT optimal cutoff score from their data could expect that 62% or more of the placement decisions made would be correct. Differentiating by course type shows that intermediate algebra courses (with the ACT mathematics score used for placement) were among the courses with the lowest median accuracy rate (62%), while English Composition II courses (with the ACT English score used for placement) had the highest (68%). Although the magnitude of the accuracy rates might be used as evidence of placement validity, one needs to compare the maximum accuracy rate at the optimal cutoff score to the accuracy rate that would result without placement—the accuracy rate that would result if all students were allowed to enroll in the course. The difference between these two values for each course represents the increase in the accuracy rate resulting from using ACT test scores for placement. For example, for college algebra, the median accuracy rate was 66%, and the median increase in accuracy rate was 13 percentage points. Thus, if all students were allowed into the course, the expected accuracy rate would be 53%.

Courses in math, social science, and natural science tended to show higher increases in accuracy rates than English courses when ACT test scores were used for placement. For English courses with sufficiently large samples, the course placement statistics were assessed for ACT English scores. English courses tend to have higher percentages of students earning a B or higher, so the accuracy rates are well above 50% without any placement measures. This leads to smaller increases in accuracy rates regardless of the placement variable (e.g., standardized tests, high school grades, locally developed placement tests, or performance assessments).

### *Success Rate*

The median success rates at the optimal cutoff score ranged from 60% in economics and intermediate algebra courses to 68% in English composition courses. This suggests that an institution using its optimal ACT cutoff score could typically expect at least 60% of the students placed in a standard course to obtain a grade of B or higher.

The median logistic regression slopes measure the strength of the relationship between ACT test scores and course success outcomes. Specifically, the slopes represent the change in the log-odds of success for each 1-point increase in the test score. For example, the log-odds of success in biology increased by 0.196 for each 1-point increase in the ACT science score. Consistent with a prior study (Allen, 2013), the slopes tended to be larger for math and natural science courses than for English and social science courses.

The optimal cutoff score for a given course varies across institutions (Allen, 2013). Variation in grading standards and course difficulty across institutions can contribute to this variation in optimal cut scores. Because results vary across institutions, institutions should collect their own course outcome data and determine their placement cutoff scores accordingly. For more details on methods for setting institution-specific cut scores, see Section 7.3.5. Though results for the ACT writing test are not shown in Table 7.14, a recent study involving 28 postsecondary institutions indicated that scores from the enhanced ACT writing test that was introduced in fall 2015 were related to the grades earned in English composition, providing validity evidence in support of using the ACT writing test to help inform course placement decisions (Radunzel, 2019). Specifically, scores from the ACT writing test were found to contribute incrementally to the prediction of English composition grades beyond ACT English score and HSGPA.

### Summary

The use of ACT scores for placement purposes increased the accuracy rate in all courses. The increases in accuracy rates were larger in math, social science, and natural science than they were in English. However, English courses tend to have higher percentages of students earning a B or higher, leading to smaller increases in accuracy rates. This phenomenon occurs regardless of the placement variable(s) used. Lastly, results varied across institutions for all the courses examined. Consequently, ACT encourages institutions to collect their own course outcome data and determine institution-specific placement cutoff scores accordingly.

**Table 7.14.** Decision-Based Validity Statistics for Course Placement Using ACT Scores (Success Criterion = B or Higher Grade)

| Course | | ACT score | No. of institutions | Median cut score* | Median logistic slope | Maximum accuracy rate | | | Increase in accuracy rate | | | Success rate | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | $Q_1$ | Med. | $Q_3$ | $Q_1$ | Med. | $Q_3$ | $Q_1$ | Med. | $Q_3$ |
| English course | Composition I | English | 215 | 18 | 0.135 | 63 | 67 | 72 | 1 | 2 | 9 | 63 | 68 | 73 |
| | Composition II | | 62 | 19 | 0.131 | 64 | 68 | 72 | 0 | 2 | 7 | 64 | 68 | 73 |
| Math course | Elementary algebra | Math | 76 | 19 | 0.244 | 60 | 64 | 68 | 6 | 13 | 23 | 58 | 63 | 68 |
| | Intermediate algebra | | 79 | 21 | 0.203 | 59 | 62 | 65 | 5 | 14 | 21 | 56 | 60 | 63 |
| | College algebra | | 134 | 22 | 0.203 | 62 | 66 | 69 | 7 | 13 | 24 | 61 | 65 | 69 |
| | Statistics/probability | | 17 | 21 | 0.184 | 61 | 65 | 68 | 3 | 10 | 22 | 59 | 61 | 68 |
| | Precalculus | | 27 | 24 | 0.184 | 63 | 66 | 69 | 1 | 8 | 22 | 61 | 65 | 69 |
| | Trigonometry | | 41 | 24 | 0.184 | 62 | 65 | 68 | 5 | 11 | 21 | 60 | 64 | 67 |
| | Calculus | | 15 | 27 | 0.146 | 61 | 67 | 69 | 1 | 9 | 18 | 62 | 67 | 69 |
| Social science course | American history | Reading | 60 | 23 | 0.114 | 61 | 63 | 66 | 4 | 11 | 19 | 59 | 61 | 64 |
| | Other history | | 30 | 23 | 0.147 | 63 | 66 | 71 | 5 | 9 | 18 | 63 | 67 | 70 |
| | Psychology | | 107 | 22 | 0.126 | 63 | 65 | 69 | 2 | 9 | 18 | 62 | 65 | 70 |
| | Sociology | | 53 | 21 | 0.118 | 63 | 65 | 68 | 1 | 5 | 14 | 63 | 66 | 68 |
| | Political science | | 33 | 22 | 0.108 | 61 | 62 | 65 | 3 | 6 | 15 | 62 | 64 | 66 |
| | Economics | | 10 | 24 | 0.111 | 60 | 63 | 65 | 3 | 16 | 25 | 58 | 60 | 64 |
| Natural science course | Biology/life sciences | Science | 108 | 23 | 0.196 | 63 | 65 | 69 | 6 | 14 | 24 | 60 | 64 | 67 |
| | General chemistry | | 55 | 26 | 0.148 | 60 | 63 | 67 | 4 | 17 | 25 | 59 | 61 | 64 |

*Note.* Placement analyses that did not yield an optimal cutoff score (i.e., the logistic function did not include a probability of .50) were not summarized in this table.

*The median cut scores reported in the table were weighted to reflect the national population of high school graduates to be consistent with the ACT College Readiness Benchmarks. Med. = median.

### *7.3.3 Incremental Validity of ACT Scores and High School Grades in Course Placement*

ACT encourages institutions to use multiple measures for placing students into college courses. Previous studies have reported that test scores and HSGPA, when used together, provide more information than either measure alone (Noble et al., 2004; Sawyer, 2010). Specifically, the use of multiple measures often results in stronger predictive relationships with course grades and increased classification accuracy. Improved classification accuracy has important implications for institutions, especially at community colleges where large percentages of students enter college academically unprepared and require remediation (Sparks & Malkus, 2013). A recent study examined the joint use of ACT scores and HSGPA for course placement at community colleges to demonstrate how using multiple measures can result in more informed placement decisions (Westrick, 2016). This study concluded that institutions can more accurately predict a student's chance of success in college courses when they use more than one measure.

Supplemental analyses using the same dataset were conducted to obtain the median accuracy rate (percentage of correct placement decisions), the median increase in accuracy rate (increment in percentage of correct placement decisions over using no predictors), and observed success rate (Table 7.15). For both English Composition I and college algebra, the joint use of ACT test scores and HSGPA resulted in the highest accuracy rates, indicating that institutions can make better placement decisions if they use both ACT test scores and HSGPA together for placement. Additional information on the methodology used in these supplemental analyses can be found in a report by Westrick and Allen (2014) that conducted similar analyses using ACT Compass scores instead of ACT scores (the ACT Compass test has since been retired).

**Table 7.15.** Median Placement Statistics for ACT Scores and HSGPA as Predictors at Community Colleges

| Course | No. of institutions | No. of students | Predictor | Median accuracy rate | Median increase in accuracy rate | Observed success rate |
|---|---|---|---|---|---|---|
| English Composition | 259 | 288,266 | ACT English | 63.3 | 4.9 | 60.6 |
| | | 256,110 | HSGPA | 66.7 | 8.3 | 61.2 |
| | | 256,100 | ACT English, HSGPA, & ACT English × HSGPA | 66.8 | 7.9 | 61.2 |
| College algebra | 182 | 132,850 | ACT Mathematics | 66.5 | 25.9 | 42.2 |
| | | 119,228 | HSGPA | 67.7 | 19.9 | 43.2 |
| | | 119,228 | ACT Mathematics, HSGPA, & ACT Mathematics × HSGPA | 68.6 | 24.5 | 43.2 |

*Note.* Success rates varied across the three analyses for each course because the datasets were slightly different (not all students had both ACT scores and HSGPA data). Observed success rates (percentages of students with a B or higher) were calculated across all institutions combined. Accuracy rates were calculated at each institution.

Sanchez (2024a) examined whether the relationships between college readiness measures and course placement decisions have changed in recent years. The study used data from 40,194 ACT-tested graduates in the 2017–2021 cohorts enrolled at 10 public 4-year institutions in one statewide-testing state. Using hierarchical logistic regression with random intercepts for institutions, the study found that higher HSGPA and higher ACT Composite scores each reduced the probability of developmental course placement across cohorts. However, the relationship between HSGPA and developmental course placement weakened across cohorts, consistent with institutions adjusting placement thresholds in response to grade inflation, while the relationship between ACT Composite score and developmental placement remained stable. The interaction between HSGPA and ACT Composite score was a significant predictor of developmental course placement, supporting the use of both measures for placement decisions. Although based on a single state, these findings are consistent with the broader argument that multiple measures are needed for course placement decisions, particularly during periods of rising grade inflation.

### *7.3.4 Differential Prediction by Student Demographic Groups in Course Placement*

A study by Allen (2016b) examined the predictive validity of using ACT scores for course placement by student demographic group. The study focused on four student demographic groups: English learners, students with disabilities, students of color (African American, Native

American, Hispanic, Native Hawaiian, students of multiple races, and students of other races, not including White and Asian), and students from low-income households (<$36,000). Specifically, the study examined how much ACT cut scores associated with a 50% chance of earning a B or higher varied by demographic group.

The results of this study are consistent with prior research showing slight overprediction for students with disabilities (Huh & Huang, 2016; Ziomek & Andrews, 1996), students of color (Lorah & Ndum, 2013; Noble et al., 1996; Sanchez, 2013; Sawyer, 1985), and students from low-income households (Lorah & Ndum, 2013; Sanchez, 2013) and slight underprediction for English learners (Mattern et al., 2008; Patterson & Mattern, 2012) when standardized test scores are used to predict individual first-year course grades and overall FYGPA. Despite some of these differences, the accuracy rates at optimal ACT cutoff scores associated with predicting first-year course success were found by Noble et al. (1996) to be somewhat comparable across gender and racial/ethnic groups. Moreover, that research also identified similar patterns of overprediction and underprediction by gender and race/ethnicity when using high school subject-area GPAs alone to predict first-year college grades. Taken together, these findings highlight the importance of using multiple measures in making course placement decisions. This statement is further substantiated by a study showing that psychosocial constructs (motivation and self-regulation) helped explain the gender gaps in first-year course outcomes that were observed after the researchers adjusted for ACT scores and for the type of college and the admissions policies of the college the student attended (Ndum et al., 2015).

### *7.3.5 Methods for Setting Cutoff Scores*

Institutions have unique placement needs that require locally developed cutoff scores rather than the median optimal cutoff scores from the studies described above. This section provides guidance concerning procedures that institutions might follow to set cutoff scores. There are multiple ways to establish cutoff scores or decision zones for placing students into different courses. The procedures for setting cutoff scores include using logistic regression and decision-based statistics, evaluating local score distributions (often with respect to institutional resources), using judgmental procedures based upon a content review of the items, and making comparisons with reference populations.

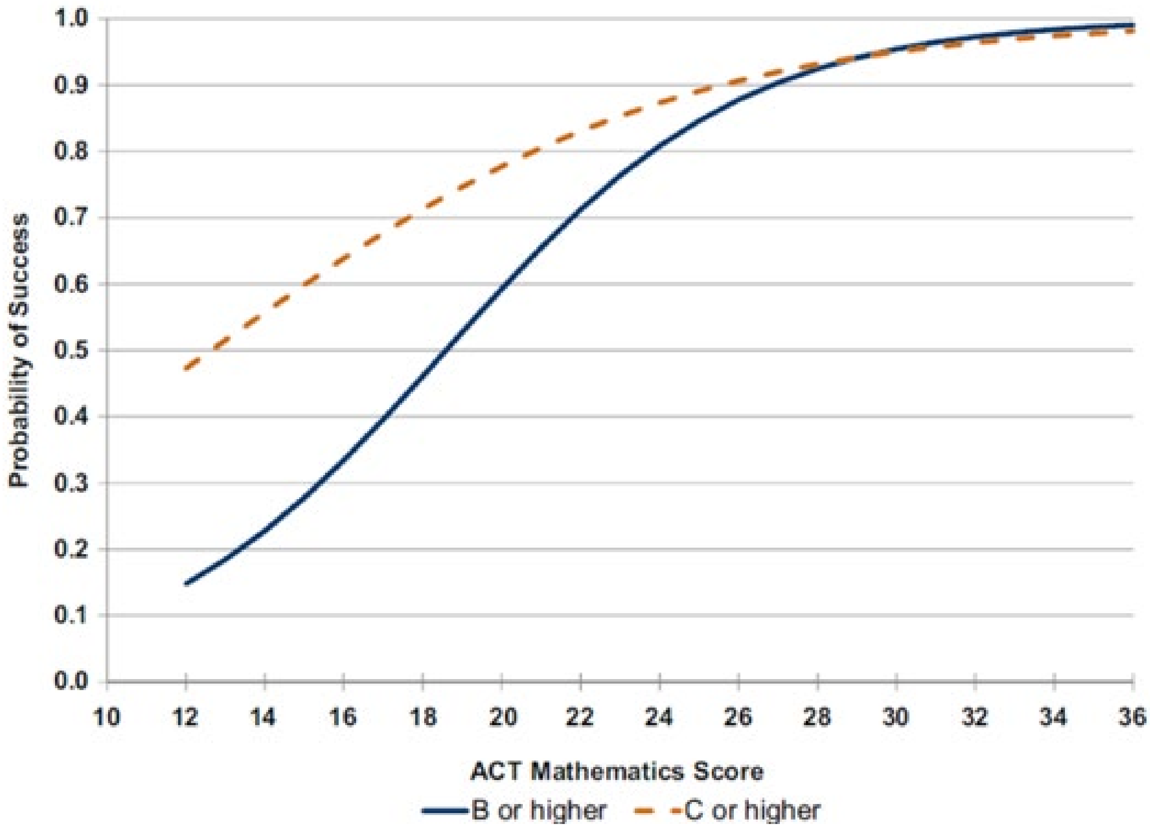
It is often advisable to interpret cutoff scores as guides rather than as rigid rules. One way to do this is to use decision zones. A decision zone is an interval around a cutoff score; students whose test scores (or other variable values) are in a decision zone are encouraged to provide more information about their academic qualifications and skill levels. For example, it might be appropriate to identify an ACT English score range of 17–20 as a placement decision zone for composition courses. Students whose scores are above 20 would be placed into composition. Those with scores below 17 would be placed into a developmental writing course that prepares them for composition. Students whose scores fall in the decision zone would be advised that their skills appear to be on the borderline of readiness for composition. Their options (which would best be discussed with an advisor) would be to enroll in a developmental course (or participate in other appropriate skill-building services) or to enroll directly in the composition course, with full awareness that most of the other students would probably have stronger prerequisite skills. To provide more information about their readiness for composition, these students could also take another test of their writing skills.

A course placement study generates the probability of success, accuracy rate, success rate, and percentage not admitted or percentage placed in a lower-level course. If a test is effective for placement, then higher test scores should correspond to higher probabilities of success. Information about students' probability of success can be used for advising individual students. It also serves as the basis for computing the group statistics used to validate tests and select cutoff scores. As an example, Table 7.16 shows the relationship between students' ACT mathematics scores and their probability of earning a B or higher and a C or higher in Mathematics 100, a course at an institution. In this course, an ACT mathematics score of 18 corresponds to a .46 probability of earning a B or higher. That is, 46 out of 100 students with an ACT mathematics score of 18 would be expected to earn a B or higher in Mathematics 100. This information is also shown graphically in Figure 7.9.

**Table 7.16.** Probability of Success in Mathematics 100 Given ACT Mathematics Score

| ACT mathematics score | Probability of success (B or higher) | Probability of success (C or higher) |
|---|---|---|
| 34 | .98 | .97 |
| 33 | .98 | .97 |
| 32 | .97 | .96 |
| 31 | .96 | .96 |
| 30 | .95 | .95 |
| 29 | .94 | .94 |
| 28 | .92 | .93 |
| 27 | .90 | .92 |
| 26 | .88 | .91 |
| 25 | .85 | .89 |
| 24 | .81 | .87 |
| 23 | .76 | .85 |
| 22 | .71 | .83 |
| 21 | .65 | .81 |
| 20 | .59 | .78 |
| 19 | .53 | .75 |
| 18 | .46 | .71 |
| 17 | .40 | .68 |
| 16 | .33 | .64 |
| 15 | .28 | .60 |
| 14 | .23 | .56 |
| 13 | .18 | .51 |
| 12 | .15 | .47 |

**Figure 7.9.** Probability of Success in Mathematics 100 Given ACT Mathematics Score




Decision-based statistics provide information about how a placement system affects groups of students. Such group-level information is important in validating and selecting cutoff scores for placement. The percentage of students who would be placed in lower-level courses is one important consideration. The availability of instructors, classrooms, and other resources affects how many students can be enrolled in either standard or lower-level courses. Moreover, if a test is effective for placement, then it should have a high estimated accuracy rate. That is, whether students are placed in a standard course or a lower-level course, the decision should be correct more often than not. Finally, using an effective placement test should also result in a high estimated success rate, which means that most students placed into a course should be successful.

Table 7.17 is provided as an example of these statistics. If an ACT mathematics cutoff score of 20 were used for placement into Mathematics 100, then about 54% of the students would be placed into a lower-level course. With respect to the success criterion of a B or higher, about 69% of all the placement decisions (into either course) would be correct; of the students placed into Mathematics 100, about 76% of them would be expected to succeed.

The optimal cutoff score is a reasonable starting point for the selection process and can be found by identifying the score that corresponds to a probability of success of about 0.50. In Tables 7.16 and 7.17, the ACT mathematics score of 19 is the cutoff score associated with at least a 50% chance of earning a B or higher and the score that would maximize the accuracy of placement into Mathematics 100 (69%) for the B or higher success criterion.

One should keep in mind, however, that the cutoff score that maximizes the accuracy rate may be associated with a success rate and a percentage of students not admitted (or placed in the lower-level course) that is not acceptable to an institution. In Table 7.17, using the optimal cutoff (ACT mathematics score of 19) would place approximately 46% of the students into the lower-level course, and with respect to the B or higher success criterion, about 73% of the students enrolled in Mathematics 100 would be successful. A lack of resources may make it impossible for an institution to place 46% of its students into lower-level courses. A solution might be to use a cutoff score of 18. This would result in an accuracy rate nearly identical to the rate associated with a score of 19 (68% versus 69%), but only 38% of students would be placed into the lower-level course. The disadvantage of lowering the cutoff score would be that the percentage of students estimated to be successful in Mathematics 100 would decrease from 73% to 69%. The institution would need to consider the consequences of selecting alternative cutoff scores as they relate to resources, as well as to institutional goals and policies.

**Table 7.17.** Decision-Based Statistics for Placement Based on ACT Mathematics Score

| ACT mathematics score | Percentage placed in lower-level course | B or higher | | C or higher | |
|---|---|---|---|---|---|
| | | Estimated accuracy rate | Estimated success rate | Estimated accuracy rate | Estimated success rate |
| 34 | 99 | 45 | 98 | 24 | 98 |
| 33 | 99 | 45 | 98 | 24 | 97 |
| 32 | 99 | 45 | 98 | 24 | 97 |
| 31 | 99 | 45 | 98 | 24 | 97 |
| 30 | 99 | 45 | 97 | 24 | 97 |
| 29 | 99 | 45 | 96 | 25 | 96 |
| 28 | 97 | 47 | 93 | 27 | 94 |
| 27 | 94 | 49 | 92 | 29 | 93 |
| 26 | 90 | 53 | 90 | 32 | 92 |
| 25 | 85 | 56 | 88 | 36 | 91 |
| 24 | 79 | 60 | 86 | 41 | 90 |
| 23 | 71 | 64 | 83 | 47 | 89 |
| 22 | 65 | 66 | 81 | 51 | 88 |
| 21 | 58 | 68 | 78 | 55 | 87 |
| 20 | 54 | 69 | 76 | 57 | 86 |
| 19 | 46 | 69 | 73 | 61 | 85 |
| 18 | 38 | 68 | 69 | 65 | 83 |
| 17 | 24 | 65 | 63 | 70 | 80 |
| 16 | 10 | 60 | 59 | 74 | 78 |
| 15 | 2 | 57 | 56 | 76 | 77 |
| 14 | 0 | 56 | 56 | 76 | 76 |
| 13 | 0 | 55 | 55 | 76 | 76 |

### Local Score Distributions

Institutional personnel are often required to establish cutoff scores based on administrative considerations (e.g., availability of instructional staff and facilities). Score distributions can be used under these conditions to provide preliminary cutoff scores. Cutoff scores based on score distributions are easy to communicate and to implement in placement systems. However, students' true abilities may be inconsistent with the selected cutoff score; that is, students who are underprepared for college may be incorrectly placed in the standard course. For more accurate decisions, ACT scores or other variables related to college or course outcomes should be used for placement decisions.

### Expert Judgment

When expert judgment is used to establish cutoff scores, institutional personnel should conduct a thorough review of the test content. Based on this review, institutions may determine that a student correctly answering a certain percentage or more of the items has demonstrated sufficient knowledge of the subject to be placed in a particular course.

There are a variety of methods for determining the cutoff score associated with the minimum skill level required for placement. (For a description of some of these methods, see Cizek & Bunch, 2006.) These methods often require content experts to judge how a "borderline" test taker (i.e., one whose knowledge and skills barely reach the decision zone) would perform on each item. Since each of these methods relies on subjective judgment, inspection of actual performance data is also recommended.

### Other Comparison Populations

Cutoff scores can also be set using the scores from the ACT national norms or Table 7.14. This is particularly helpful when local normative data are not available. The normative distribution would be used in a manner similar to that described above for local score distributions. A student taking a specific test would be placed in a standard course if he or she scored at or above the scale score corresponding to a predetermined percentile rank in the score distribution of the reference population. Users should note that local distributions of ACT scores and grades may differ markedly from national distributions. Therefore, cutoff scores derived from national data should be validated and later adjusted as warranted when local data become available.

### Monitoring Cutoff Scores

Once an institution selects a procedure and establishes a cutoff score, it is essential for the institution to continually monitor the effectiveness of the cutoff score. Experience may suggest adjusting established cutoff scores.

## 7.4 Using ACT Scores to Evaluate Students' Likelihood of College Success in the First Year and Beyond

Sections 7.2 and 7.3 summarized the results of various studies that examined the relationships between ACT scores and first-year course grades as they relate to admissions and placement decisions. This section describes studies illustrating the relationship between college readiness as measured by the ACT and students' success as indicated by additional outcomes from the first year of college and beyond. The information presented in this section may be considered further evidence supporting the use of ACT scores for making college admissions decisions. The first subsection focuses on relating ACT Benchmark attainment to first-year outcomes that include college enrollment, first-year college grades, and college retention. The second subsection focuses on relating ACT scores to ACT Collegiate Assessment of Academic Proficiency (CAAP) scores earned by students during their second year of college. The third and fourth subsections focus on relating ACT scores to longer-term outcomes that include cumulative college GPA at graduation and degree attainment. The fifth subsection focuses on relating the ACT STEM score to students' chances of persisting and completing a college degree in a STEM-related field. The sixth subsection describes predictive validity for students testing with accommodations. The final subsection discusses statistical relationships between ACT scores and cumulative first-year hours earned.

### *7.4.1 Statistical Relationships Between College Readiness and First-Year College Success*

This section provides estimates of students' chances of college success as measured by several different first-year outcomes, examined by ACT College Readiness Benchmark attainment in individual subject areas as well as by the number of Benchmarks met (see Chapter 5 or Allen, 2013, for a description of the Benchmarks). The results presented here, based on more recent freshman cohorts, update some findings from an earlier study conducted by ACT (2010).

#### Data and Method

College outcomes included enrollment into any college the fall following high school graduation, earning a B or higher in first-year college courses, achieving a FYGPA of 3.0 or higher, and remaining enrolled at the initial institution in the second year. College readiness was measured by ACT College Readiness Benchmark attainment.

College enrollment rates were based on approximately 1.9 million high school students who took the ACT and indicated that they would graduate from high school in 2015. Colleges included both 2-year and 4-year institutions. College retention rates were based on approximately 1.3 million ACT-tested students from the 2015 graduating class who enrolled in a postsecondary institution the fall following high school graduation according to the National Student Clearinghouse database. More than 2,800 colleges were included. Data for FYGPA included approximately 430,000 ACT-tested students from nearly 300 postsecondary institutions who participated in research services offered by ACT. Approximately 125,000 students were included in the analysis for English Composition I; 31,000 for English Composition II; 20,000 for intermediate algebra; 69,000 for college algebra; 5,000 for precalculus/finite math; 18,000 for calculus; 41,000 for American history; 77,000 for psychology; 32,000 for biology; and 31,000 for chemistry. For all outcomes except college enrollment, the study used hierarchical logistic regression models to estimate students' chances of success as a function of Benchmark attainment or the number of Benchmarks met while statistically controlling for the institution attended. Random intercept models were estimated. For college enrollment, observed rates were calculated.

### Results

Students who met the ACT College Readiness Benchmarks were more likely than those who did not to

- enroll in college the fall following high school graduation (by 23 to 29 percentage points; Figure 7.10),
- earn a B or higher in first-year college courses (by 18 to 27 percentage points; Figure 7.11),
- achieve a FYGPA of 3.0 or higher (by 23 to 27 percentage points; Figure 7.12), and
- remain enrolled at the same institution in the second year (by 6 to 9 percentage points; Figure 7.13).

Moreover, as the number of Benchmarks attained increased, students' likelihood of success also increased for each of the first-year outcomes examined (Table 7.18). For example, students' chances of enrolling in college increased from 45% for those who met none of the Benchmarks to 83% for those who met all four Benchmarks.

**Figure 7.10.** College Enrollment Rates by ACT College Readiness Benchmark Attainment

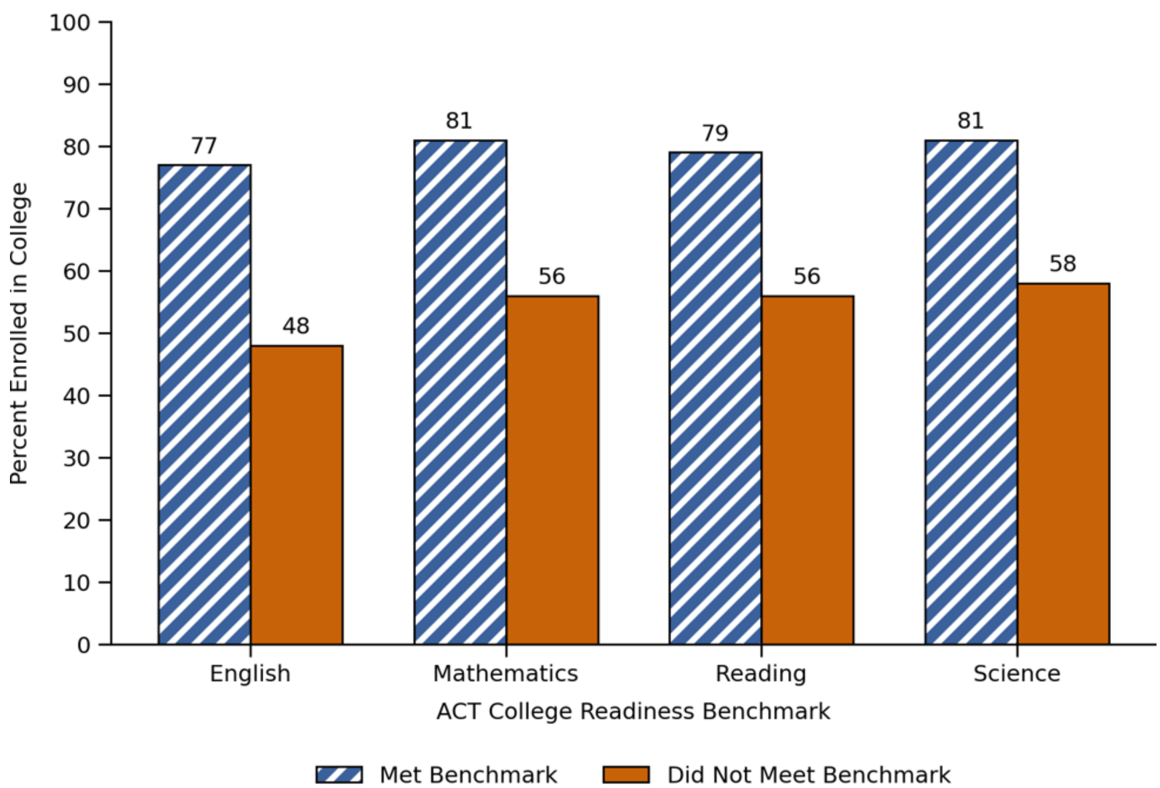


**Figure 7.11.** Students' Chances of Earning a B or Higher in First-Year College Courses by ACT College Readiness Benchmark Attainment at a Typical Institution

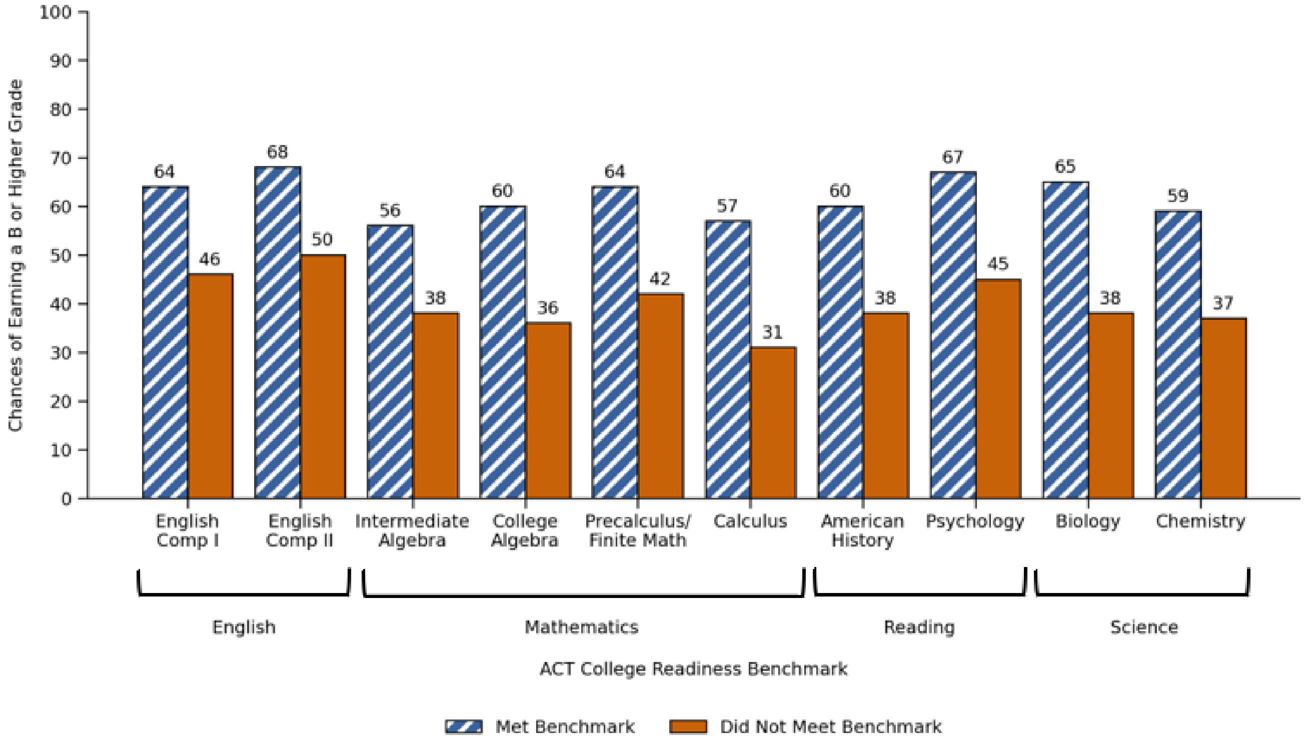


**Figure 7.12.** Students' Chances of Achieving a 3.0 or Higher FYGPA by ACT College Readiness Benchmark Attainment at a Typical Institution

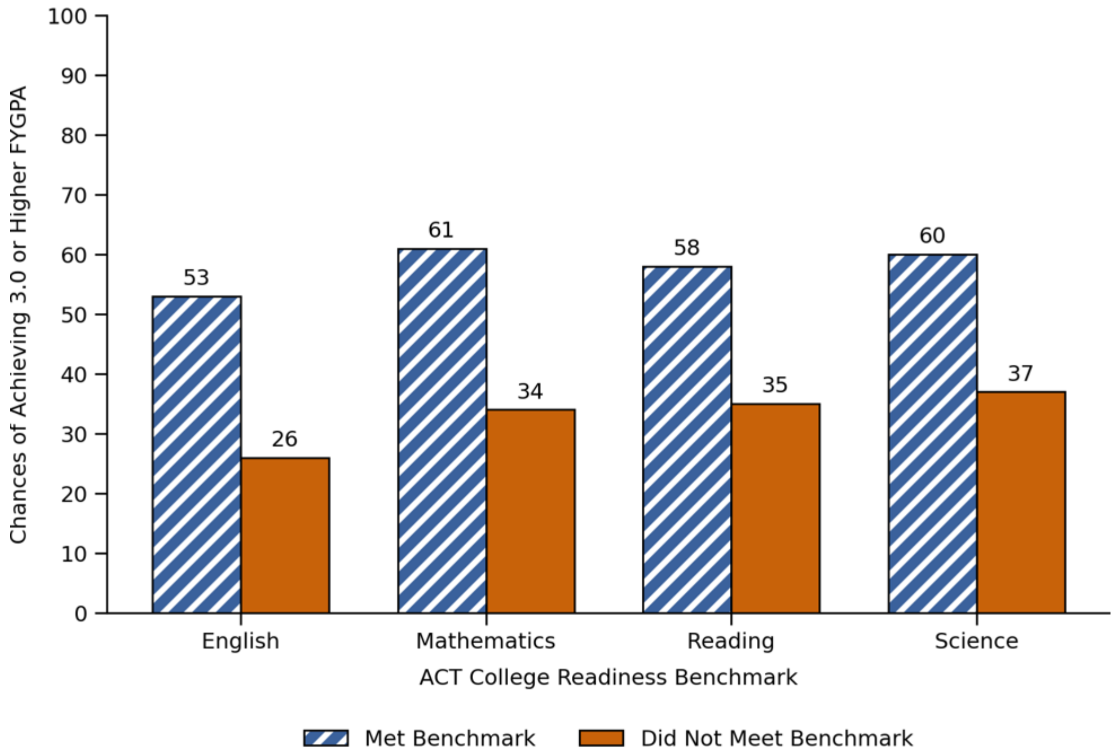



**Figure 7.13.** Students' Chances of Remaining Enrolled at the Initial Institution in Year 2 by ACT College Readiness Benchmark Attainment

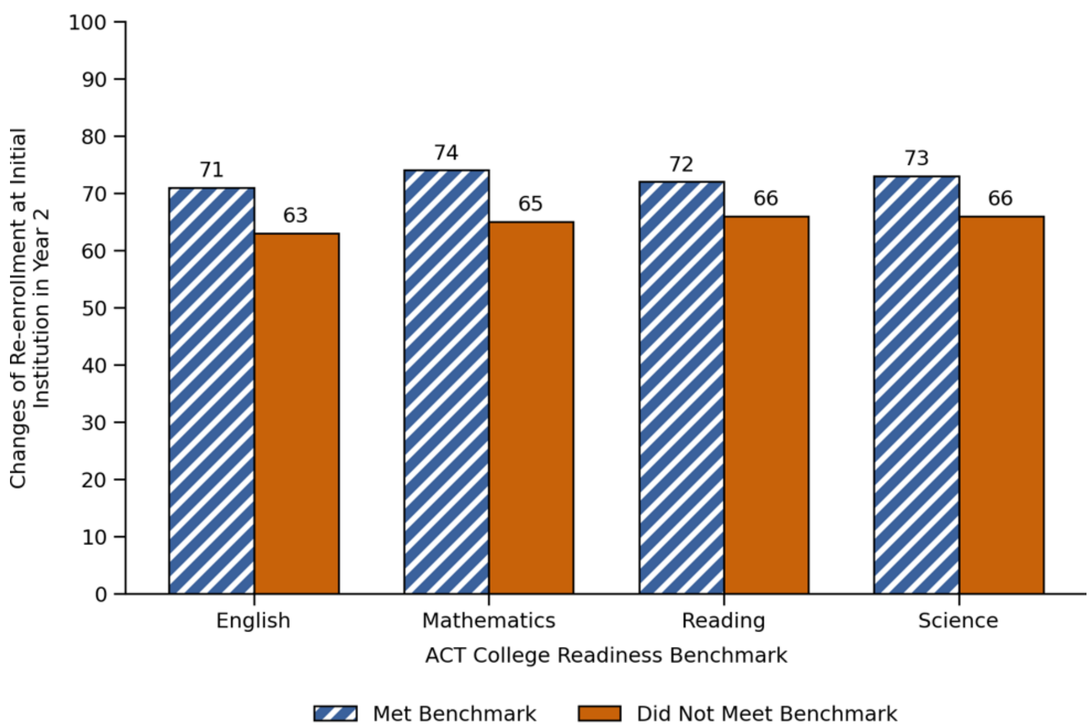


**Table 7.18.** Percent Chance of First-Year College Outcomes by Number of ACT College Readiness Benchmarks Met

| Outcome | | Number of ACT Benchmarks met | | | | |
|---|---|---|---|---|---|---|
| | | 0 | 1 | 2 | 3 | 4 |
| **Enrollment** | | 45 | 66 | 73 | 78 | 83 |
| **B or higher in** | English Composition I | 45 | 54 | 61 | 68 | 75 |
| | English Composition II | 51 | 58 | 65 | 71 | 76 |
| | Intermediate algebra | 33 | 39 | 46 | 52 | 58 |
| | College algebra | 29 | 37 | 46 | 55 | 64 |
| | Precalculus/finite math | 38 | 45 | 52 | 60 | 66 |
| | Calculus | 25 | 33 | 42 | 51 | 61 |
| | American history | 29 | 40 | 51 | 62 | 72 |
| | Psychology | 35 | 47 | 58 | 69 | 79 |
| | Biology | 22 | 33 | 47 | 62 | 75 |
| | Chemistry | 21 | 30 | 41 | 53 | 65 |
| **FYGPA of 3.0 or higher** | | 22 | 33 | 45 | 57 | 69 |
| **Retention** | | 62 | 66 | 69 | 73 | 76 |

### Summary

The ACT College Readiness Benchmarks are good indicators of whether students have acquired the knowledge and skills needed to succeed in first-year college courses. The results from the current analyses also show that students who are better prepared academically for college (as indicated by meeting the Benchmarks) are more likely than less-prepared students to immediately enroll in college and tend to be more successful during their first year of college and to remain enrolled at their initial institution in the second year.

## *7.4.2 Statistical Relationships Between ACT and ACT CAAP Scores*

The previous section showed that students who are better prepared academically, as measured by the Benchmarks, are more likely to succeed during their first year of college than less prepared students. To better illustrate the relationship between college readiness and student academic success into the second year of college, this section describes the relationships between ACT scores, Benchmark attainment, and ACT Collegiate Assessment of Academic Proficiency (CAAP) scores for second-year college students. ACT CAAP was a standardized assessment program used by colleges and universities to evaluate student learning outcomes in their general education programs. ACT CAAP offered six independent test modules: reading, science, critical thinking, mathematics, writing skills, and essay writing (ACT, 2015a). The ACT CAAP assessment was taken by students between the academic years 2008–2009 and 2014–2015.

### Data and Method

The sample included more than 16,000 college students who took CAAP during the spring term of their second year and the ACT test in high school during their junior or senior year. Because of the modular nature of CAAP, not all students had all CAAP scores. The results were based on 11,221 students for English/writing skills, 11,892 students for mathematics, 10,574 students for reading, and 9,005 students for science. Self-reported cumulative college GPAs at the time of CAAP testing were also available as an indicator of college achievement.

College readiness was measured by ACT College Readiness Benchmark attainment (see Chapter 5 or Allen, 2013, for a description of the Benchmarks). Descriptive statistics including means, standard deviations, percentages, and correlations were used to examine how ACT scores or Benchmark attainment related to CAAP scores and cumulative college GPA in the second year.

### Results

ACT scores were strongly correlated with CAAP performance (Table 7.19). In addition, students meeting the ACT College Readiness Benchmarks in high school had higher average CAAP scores than students not meeting the Benchmarks (Table 7.20). This pattern was observed in all four content areas. The difference in average CAAP scores was as much as 6.6 points (on a 40–80 score scale). Moreover, as shown in Figure 7.14, students who met the Benchmarks in high school were more likely to have a cumulative college GPA greater than 3.0 in their second year of college.

### Summary

These findings suggest that the ACT College Readiness Benchmarks can assist in determining who will succeed in college, even into the second year.

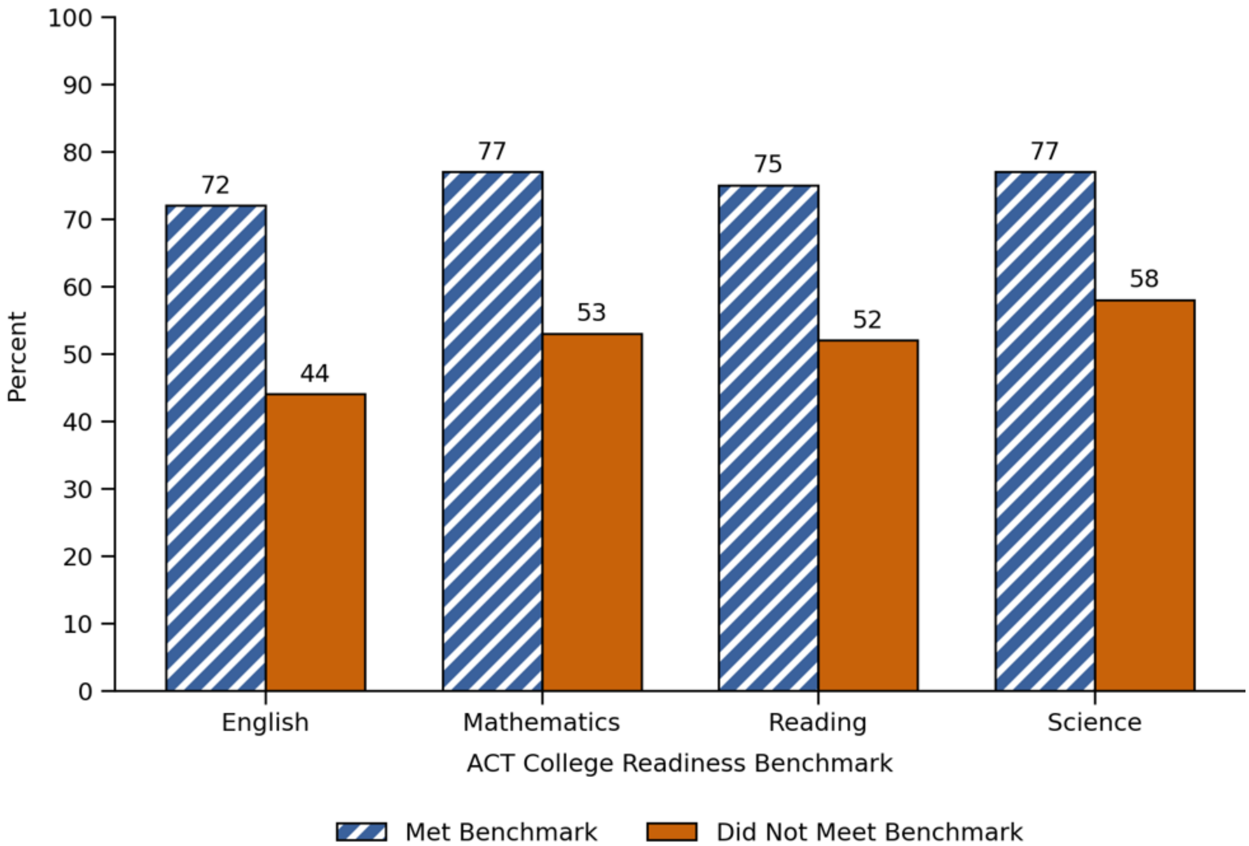
**Table 7.19.** ACT/CAAP Test Score Correlations

| English/writing | Mathematics/math | Reading/reading | Science/science |
|---|---|---|---|
| 0.77 | 0.73 | 0.70 | 0.67 |

**Table 7.20.** Average ACT CAAP Test Score by ACT Benchmark Attainment

| ACT/CAAP content area | | ACT Benchmark attainment: Met | ACT Benchmark attainment: Not met |
|---|---|---|---|
| English/writing | Mean (*SD*) | 65.0 (4.0) | 58.4 (3.5) |
| | Number of students | 8,418 | 2,803 |
| Mathematics/math | Mean (*SD*) | 60.7 (3.5) | 55.9 (3.0) |
| | Number of students | 5,145 | 6,747 |
| Reading/reading | Mean (*SD*) | 64.3 (4.6) | 58.1 (4.2) |
| | Number of students | 5,199 | 5,375 |
| Science/science | Mean (*SD*) | 63.8 (3.9) | 58.5 (3.8) |
| | Number of students | 3,514 | 5,491 |

**Figure 7.14.** Percentages Earning a Cumulative College GPA Greater Than 3.0 by ACT College Readiness Benchmark Attainment for Students Taking ACT CAAP During Sophomore Year and the ACT in High School

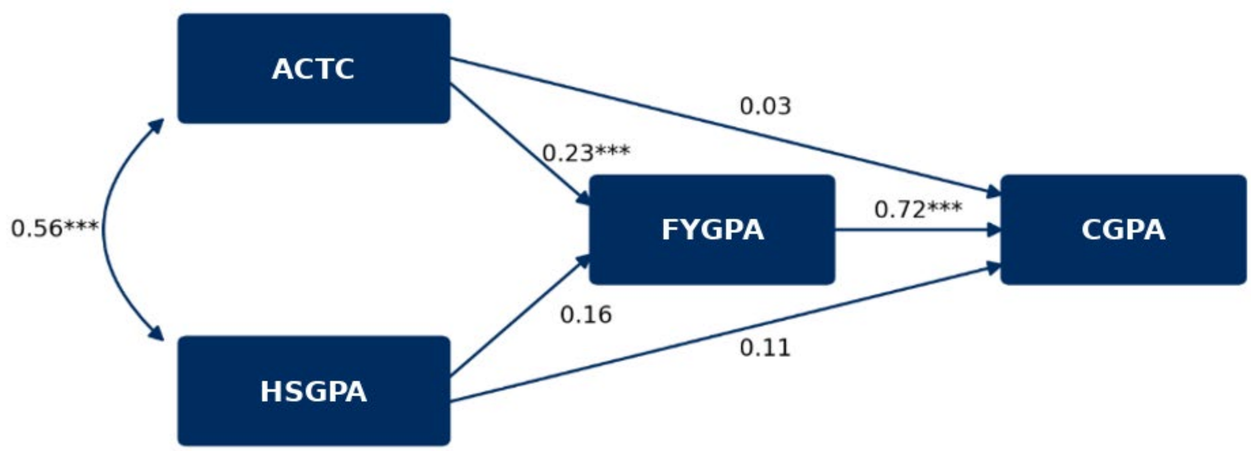



### *7.4.3 Statistical Relationships Between ACT Scores and Cumulative College GPAs*

A study by Tracey and Robbins (2006) used hierarchical linear modeling to examine the relationships between performance on the ACT and cumulative college GPA across time while accounting for the nesting of students within colleges. The findings from this study indicated that ACT performance predicted cumulative college GPA across time. The researchers also examined how congruence measures between students' interests (as measured by the ACT Interest Inventory) and college major choice related to college performance. For more details, see the full study (Tracey & Robbins, 2006).

Sanchez (2024d) used path analysis with cluster-robust standard errors to examine how HSGPA and ACT Composite score relate to cumulative GPA at graduation (CGPA), with first-year college GPA (FYGPA) modeled as a mediator. The analytic sample consisted of 7,516 students who earned a bachelor's degree within 6 years at a public 4-year institution in one statewide-testing state. As shown in Figure 7.15, ACT Composite score predicted FYGPA ($\beta = 0.23$, $p < .01$), and FYGPA in turn predicted CGPA ($\beta = 0.72$, $p < .01$), while the direct path from ACT Composite score to CGPA was small and nonsignificant ($\beta = 0.03$). FYGPA therefore fully mediated the relationship between ACT Composite score and CGPA. The corresponding paths for HSGPA were smaller in magnitude (HSGPA to FYGPA, β = 0.16; HSGPA to CGPA, β = 0.11), and HSGPA and ACT Composite score were themselves correlated ($r = .56$, $p < .01$). Beyond the mediation model, Sanchez reported that adding ACT Composite score to a model containing HSGPA improved fit relative to either predictor alone, consistent with prior work showing that the two measures contribute unique variance to the prediction of long-term college outcomes. Because they are based on a single statewide-testing state and are further restricted to students who persisted to a bachelor's degree, these findings should be interpreted with caution and read as corroborating, rather than independently establishing, the broader multistate evidence.

**Figure 7.15.** Path Analysis of the Effects of ACT Composite Score and HSGPA on CGPA With FYGPA as a Mediator




*Note.* ***$p < 0.01$

### *7.4.4 Statistical Relationships Between ACT Scores and Degree Completion*

Long-term student success is an important goal for students and postsecondary institutions. A study by Radunzel and Noble (2012) examined the relationships between performance on the ACT and degree completion at both 2- and 4-year institutions. Such information might be useful for early identification of students who might benefit from additional academic and student support services upon entering college.

#### Data and Method

Data for this study included approximately 194,000 ACT-tested students who enrolled in college as first-time entering students in fall 2000 through 2006. Approximately 126,000 students who began at one of 61 four-year institutions were tracked for at least 6 years, and nearly 68,000 students who began at one of 43 two-year institutions were tracked for at least 3 years. The outcomes were bachelor's degree completion at the initial institution within 6 years for students beginning at 4-year institutions and associate's degree completion at the initial institution within

3 years for students beginning at 2-year institutions. Because many students beginning at a 2-year institution transfer to a 4-year institution without earning an associate's degree (Radunzel, 2012), associate's degree completion or transfer to an in-state 4-year institution within 3 years was also evaluated for students beginning at 2-year institutions. The latter outcome was evaluated for a subset of the 2-year data from two state systems where students could be tracked across both 2- and 4-year institutions. Hierarchical logistic regression models were used to estimate institution-specific probabilities of degree completion based on ACT scores alone and in combination with self-reported HSGPAs. The accuracy rates and increases in accuracy rates over not using the predictor were calculated at the predictor value(s) associated with a 50% chance of degree completion (for more details on these decision-based statistics, see Section 7.2.1).

### Results

As shown in Figure 7.16, as ACT Composite score increased, students' chances of completing a degree increased for both 2- and 4-year students. Additionally, as ACT Composite score increased, 2-year students' chances of completing an associate's degree or transferring to a 4-year institution increased. As an example of the increase for those beginning at a 4-year institution, students' chances of completing a bachelor's degree in 6 years was 41% for those with an ACT Composite score of 20 and 67% for those with an ACT Composite score of 30. Higher values of HSGPA were also associated with increased chances of degree completion.

The typical maximum accuracy rate and increase in accuracy rate across institutions associated with using ACT Composite score to predict bachelor's degree completion within 6 years were 64% and 24%, respectively. Similar rates were associated with using HSGPA alone (65% and 23%). In comparison, the typical maximum accuracy rate associated with using both predictors jointly was 2 to 3 percentage points higher than those based on the single-predictor models.

Figure 7.17 provides the estimated probabilities of completing a bachelor's degree within 6 years associated with different values of HSGPA and ACT Composite score. The figure illustrates the incremental usefulness of ACT scores beyond HSGPA for predicting who is likely to complete a degree. As both HSGPA and ACT Composite score increased, probabilities of success also increased. The ACT Composite score differential was larger for students with higher HSGPAs than for those with lower HSGPAs. The same was true for the HSGPA differential in a comparison of students with higher and lower ACT Composite scores.

### Summary

Both ACT Composite score and HSGPA were effective for predicting long-term college success at 2- and 4-year institutions. Other outcomes examined in the study included progress to degree completion (based on cumulative hours earned) and cumulative GPA at degree completion. Across the outcomes, ACT test scores increased prediction accuracy over that for HSGPA alone. The study also indicated that ACT Composite scores and HSGPA were primarily indirectly related to subsequent college outcomes through FYGPA. For additional information on this study, see the full report (Radunzel & Noble, 2012).

**Figure 7.16.** Probability of Degree Completion Based on ACT Composite Score

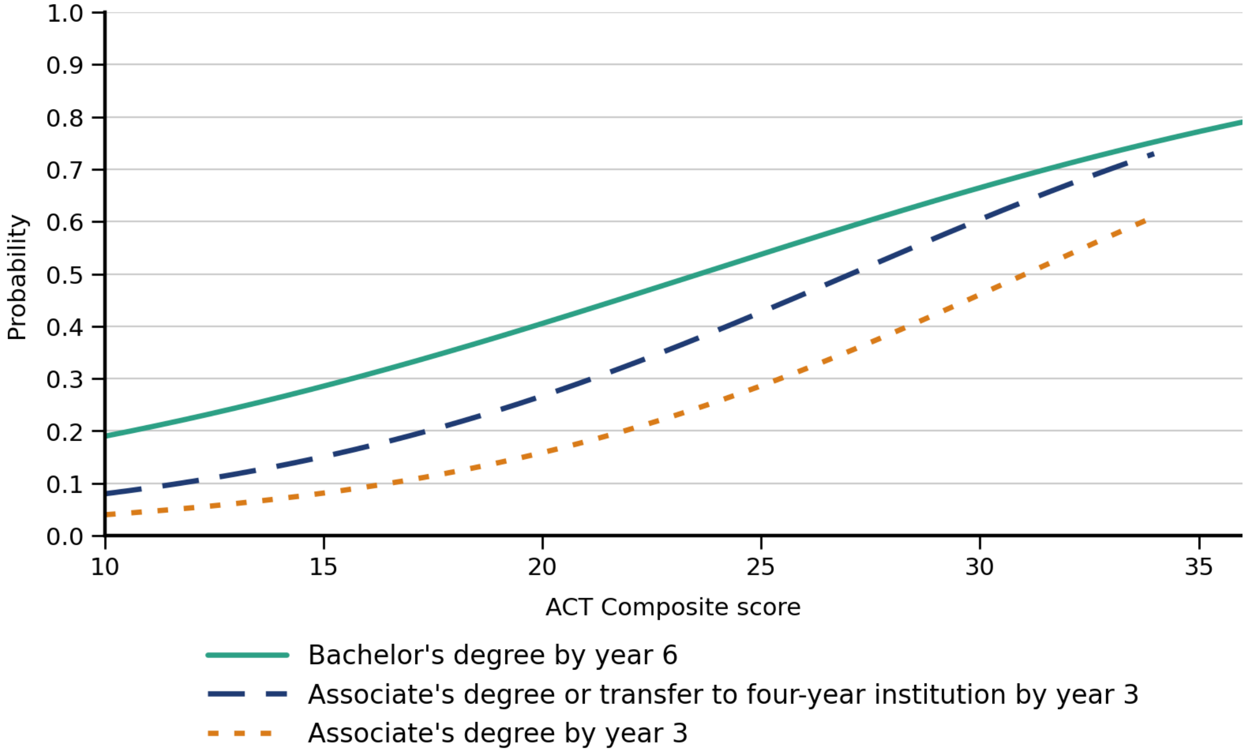
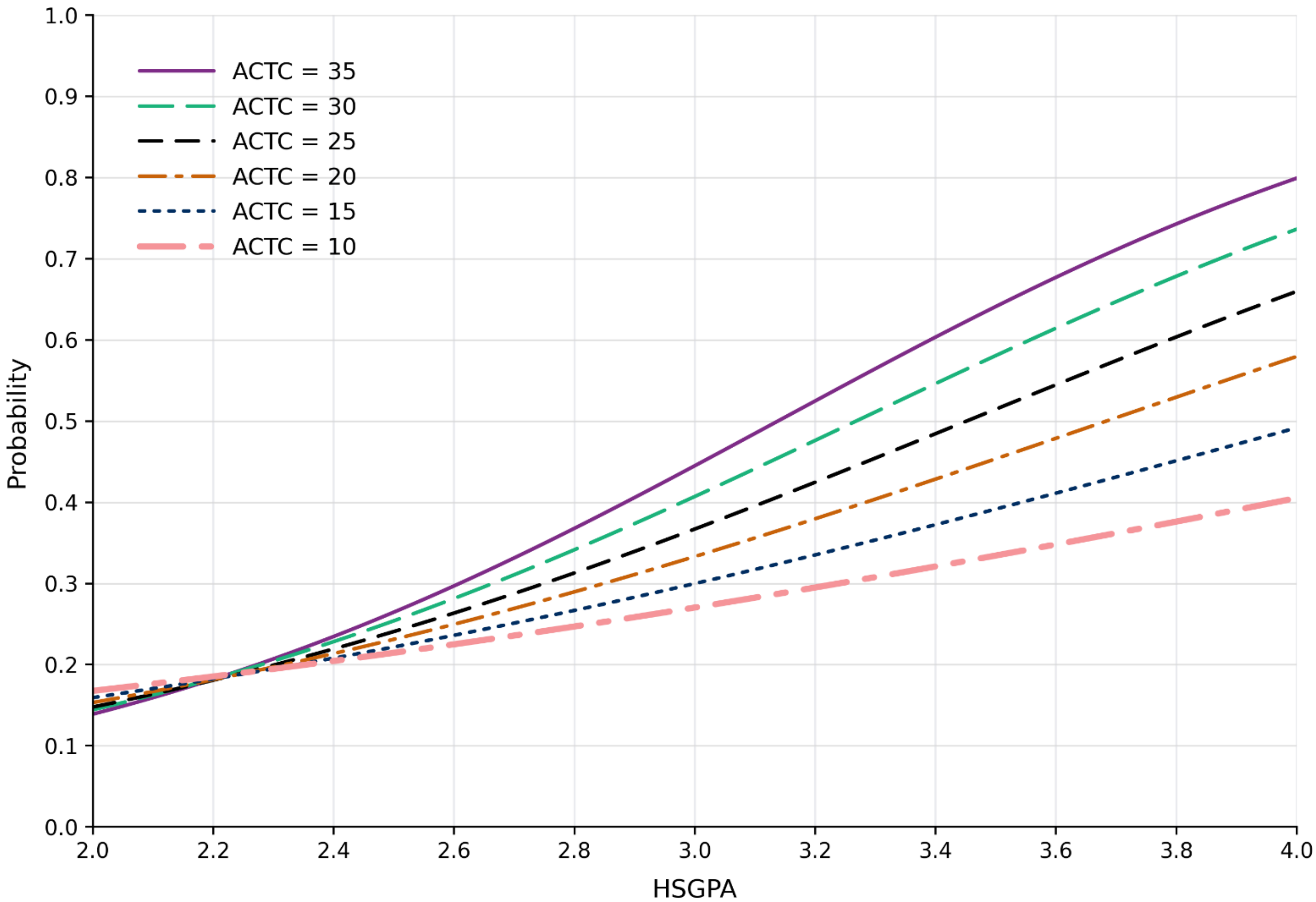
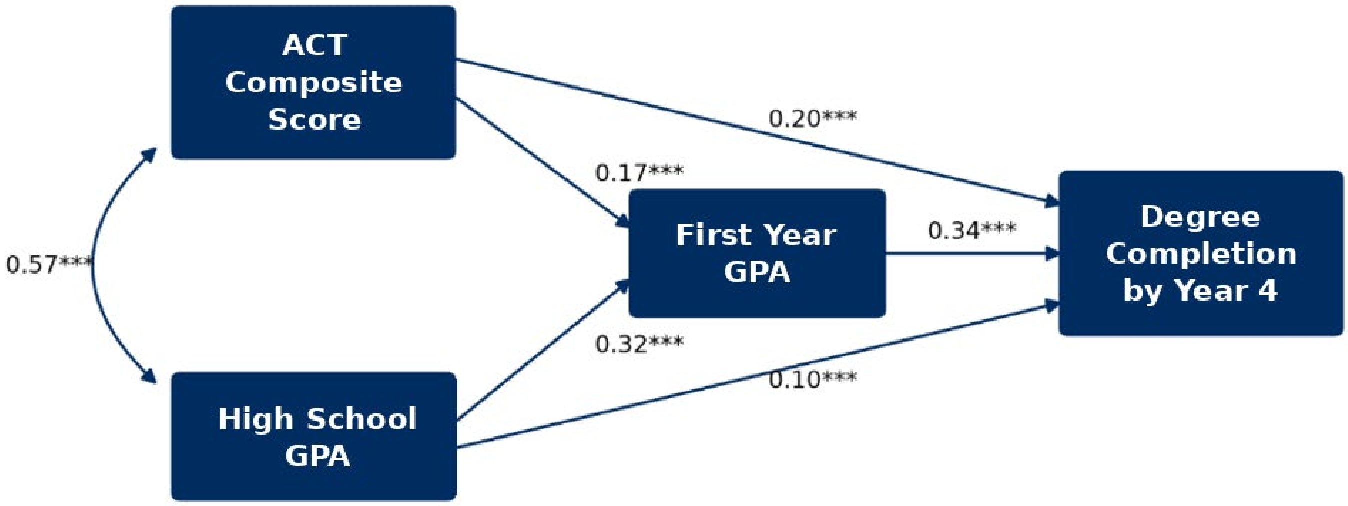


**Figure 7.17.** Probability of Bachelor's Degree Completion Within 6 Years by HSGPA and ACT Composite Score




Loria and Sanchez (2024) extended this work by formally examining FYGPA as a mediator of the relationship between college readiness measures and degree completion. Using path analysis with cluster-robust standard errors and hierarchical logistic regression, the study examined 4- and 6-year bachelor's degree completion for 19,613 first-time fall 2017 freshmen at 32 public 4-year institutions in one statewide-testing state. Both HSGPA and ACT Composite score had significant direct and indirect effects on degree completion through FYGPA; ACT Composite score had the stronger direct effect on completion, while HSGPA had the stronger indirect effect through FYGPA. A significant interaction between HSGPA and ACT Composite score was found for 4-year graduation, with ACT Composite score more strongly associated with degree completion at higher HSGPA levels. Figure 7.18 summarizes the direct and indirect effects. These results are consistent with prior multistate evidence that HSGPA and ACT Composite score contribute unique information about student readiness and long-term postsecondary success. Because they come from a single statewide-testing state, they corroborate rather than independently establish that conclusion.

**Figure 7.18.** Path Analysis of Effect of ACT Composite Score and HSGPA on Probability of Degree Completion by Year 4 With FYGPA as a Mediator




*Note.* *p < 0.1; **p < 0.05; ***p < 0.01

### *7.4.5 Statistical Relationships Between ACT STEM Scores and Students' Chances of Succeeding in a STEM-Related Major*

A study by Radunzel et al. (2015) examined the ACT STEM score in relation to the likelihood of success as measured by a variety of STEM-related college outcomes: cumulative GPA over time, persistence in a STEM major, and ultimately completing a STEM degree. The results from the study illustrated that predicting student success in STEM-related fields is a valid use of the ACT STEM score.

A study by Radunzel et al. (2016) suggested that the positive relationship between ACT STEM scores and students' chances of succeeding in a STEM major held even after the researchers statistically controlled for other student characteristics, such as high school coursework taken and grades earned, vocational interests, and demographic characteristics. This finding was consistent with a growing body of literature indicating that educational success is a product not just of academic skills and knowledge but also of noncognitive factors such as motivation, academic goals, and academic self-efficacy (Mattern et al., 2014).

Sanchez (2025d) examined the predictive validity of ACT STEM scores and HSGPA for FYGPA using data from 2,691 students in the 2022 graduating class who enrolled at a public postsecondary institution and declared a STEM major in one statewide-testing state. Using hierarchical linear models with students nested within colleges, the study found that both HSGPA and ACT STEM scores were significant independent predictors of FYGPA, with HSGPA slightly stronger as a single predictor. A significant interaction between HSGPA and ACT STEM scores indicated that students who scored highly on both measures earned substantially higher FYGPA than would be expected from either measure alone. Although limited to a single state, these findings extend the evidence for ACT STEM score validity to the specific population of declared STEM majors and support the joint use of HSGPA and ACT scores for identifying STEM-ready students.

### *7.4.6 Predictive Validity for Students Testing With Accommodations*

Moore and Schnieders (2025) explored the predictive validity of ACT Composite scores and HSGPA for first-year college GPA (FYGPA) and degree completion for disabled students who tested with accommodations. The sample included 143,768 students who graduated high school between the academic years of 2008 and 2016 and went on to enroll full time at 1 of 22 public 2-year institutions or 10 public 4-year institutions in a single southern state. Among this sample, 2,659 (2%) tested with accommodations and 141,109 (98%) did not. At 4-year institutions, the correlation between ACT Composite score and FYGPA for students testing with accommodations was 0.37, which exceeded the HSGPA and FYGPA correlation of 0.27. This is an inversion of the typical result found for students who test without accommodations, where HSGPA is usually the stronger correlated variable. The authors note that this finding suggests that the ACT may be capturing some aspects of college readiness for accommodated students that HSGPA is not. Further significant interaction effects were also found that suggest that the relationship between HSGPA and college outcomes was flatter for students who tested with accommodations than for students who did not, meaning HSGPA discriminates less well among students with disabilities who test with accommodations. The adjusted $R^2$ for predicting FYGPA at 4-year institutions was 0.22 for ACT Composite score alone, 0.25 for HSGPA alone, 0.31 for ACT Composite and HSGPA together, and 0.33 when demographics were included in the Composite/HSGPA model. The Akaike information criterion and Bayesian information criterion values for the logistic regression models that predicted degree completion consistently showed that the best fit was obtained from models that included both ACT Composite score and HSGPA along with demographic variables. At 4-year institutions, the degree completion rates were lower for students who tested with accommodations than for those who tested without: 21% versus 30% for 4-year graduation and 45% versus 51% for 6-year graduation. Additionally, the residual analysis conducted suggests that the models tend to overpredict outcomes for students who test with accommodations.

### *7.4.7 Statistical Relationships Between ACT Scores and Cumulative First-Year Hours Earned*

Sanchez (2024b) examined the predictive validity of HSGPA and ACT Composite scores for cumulative first-year hours earned across the 2018–2022 cohorts using data from 61,110 students at 32 two- and four-year institutions in one statewide-testing state. Three hierarchical linear models were fit with random slopes and intercepts for institutions: HSGPA alone, ACT Composite alone, and both measures combined.

HSGPA-based predictions of first-year hours earned declined across the cohort years, consistent with eroding HSGPA validity, while ACT Composite–based predictions remained stable across cohorts. The combined model reduced but did not eliminate this decline. Over the same period, mean HSGPA in the sample rose from 3.44 in 2017 to 3.59 in 2021, while mean ACT Composite scores declined from 22.5 to 21.9. Although based on a single state, these results suggest that grade inflation may be eroding the predictive validity of HSGPA for short-term postsecondary outcomes, while ACT Composite scores contribute stable predictive information across cohorts.

## 7.5 Using ACT Scores to Assist With Program Evaluation

The ACT test was developed to measure academic skills and knowledge that are learned in high school and are necessary for academic success in the first year of college. Validity evidence for using the ACT as a measure of educational achievement is documented at the beginning of this chapter. Since the ACT measures important educational outcomes, it might be considered for use in evaluating the effectiveness of school programs.

Before using the ACT in program evaluation, a school should conduct a content review to determine the extent to which the test represents important outcomes the school wishes to measure. If there is a content match between the ACT and important local educational outcomes, the ACT may be considered as one component of a program evaluation system. ACT scores should not be relied on exclusively as evidence of program effectiveness. Rather, ACT scores should be considered with other indicators of program effectiveness routinely collected by schools.

Several cautions must be kept in mind when using the ACT for program evaluation. ACT-tested students may not represent all students enrolled in the school; that is, expectations of and conclusions drawn about a select group of students who take the ACT may differ from those concerning a larger group of college-bound students or those of the high school graduating class as a whole. In cases where the school administers the ACT to all their juniors or seniors through statewide or districtwide testing programs, this is less of a concern. Additionally, a school's average ACT scores can fluctuate from year to year, as evidenced in a study by Sawyer (2013b). In that report, Sawyer described simple ways for school officials to better understand whether yearly changes or trends over time in average ACT scores for their school were unambiguously positive or negative instead of plausibly due to chance. Finally, without some measure of student achievement earlier in high school, judgments about educational development and achievement during high school may be misleading. This issue can be addressed by using the ACT in conjunction with PreACT, ACT's 10th-grade assessment program (ACT, 2025b), or PreACT 8/9 (ACT, 2025a).

### *7.5.1 Using ACT Scores as Outcomes for Program Evaluation*

ACT scores can be used in various ways for program evaluation. A school could establish expected levels of educational achievement based on ACT scores or Benchmark attainment for individual students, for the entire group of tested students, or for groups of students defined by common academic interests, high school coursework, or some other characteristic. Expected and actual levels of educational achievement could then be compared to evaluate program effectiveness.

In establishing expected levels of achievement for groups of students, several factors need to be considered, including the availability of resources both inside and outside the school, the social climate of the school, the characteristics of the students from the school who take the ACT, and the level of students' academic preparedness upon entering the school.

One way to determine expected levels of educational achievement is by estimating them with growth models that use prior measures of achievement from earlier grades. Section 8.4 provides additional discussion of how ACT-based growth models may be used to support program evaluation and school effectiveness analyses, along with cautions related to model choice, representativeness, and statistical uncertainty.

Allen (2026) explored the growth between PreACT and ACT for 467,516 students in the 2025 ACT-tested graduating class who took a PreACT assessment (PreACT 8/9, PreACT, or PreACT Secure) before the ACT. The main Composite gain score was 1.64 points, and the subject area gains ranged from 0.54 points in reading to 2.67 points in English. Composite score gains varied across race/ethnicity, with Native American students having the lowest gain (0.93 points) and Asian students having the highest gain (2.72 points). These differences narrowed after they were adjusted for high school grades and course rigor, school characteristics, and the number of months between tests. Gain scores rose with course rigor decile as well as family income. Across the 32 states examined, the unadjusted mean ACT score was 20.86 for PreACT-tested students and 19.86 for students who had not taken PreACT. The College Readiness Benchmark attainment gaps between the two groups ranged from 3.1% for the STEM Benchmark to 8.7% for the English Benchmark. After statistical adjustments were made, the adjusted Benchmark attainment gaps ranged from 2.0% to 5.3%. PreACT-tested students scored higher on the ACT and met the Benchmarks at higher rates than students who did not take PreACT; these differences shrank but remained after statistical adjustment.

Allen et al. (2023) provided six evidence-backed benefits of school-day ACT testing across all 50 states and the District of Columbia for the 2021–2022 academic school year. More than 1.3 million students took the school-day ACT. Among the schools that adopted school-day testing, the tested population shifted toward students who were underserved, with the percentage of low-income students rising from 43% before school-day testing to 51% after school-day testing. The percentage of Hispanic students rose from 10% to 14%, and the percentage of Black / African American students rose from 15% to 16%. Among these schools, the number of students meeting the Benchmarks increased by 33% in English, 25% in mathematics, 29% in reading, and 25% in science. The largest increases were seen for Hispanic students (+60%), students from underrepresented racial/ethnic groups (+47%), and low-income students (+44%). The number of students that opted in to ACT's Recruit Me service also increased 56% overall, with the largest increases for Hispanic students (+109%) and low-income students (+78%). Additional research shows that college enrollment increases of 0.3 to 10 percentage points were attributable to statewide mandatory college entrance exam testing.

Allen and Dingler (2023) explored some of the benefits of school-day senior retesting using data from 168,284 students at more than 1,300 schools across 40 states. This included only schools where at least 10% of the spring Grade 11 testers went on to retest in the fall of Grade 12. School-day senior retesters were more demographically diverse than students who retested on their own on national test dates. The school-day testing group was 17% Black / African

American (compared to 11% of national retesters), 14% Hispanic (compared to 11%), and 35% from families earning less than $50,000 (compared to 15%). Furthermore, 15% received English learner services (compared to 6% nationally). After the fall retest, the number of students reaching an ACT Composite score of 21, which can qualify students for scholarships in certain states, rose 18%, thereby adding 8,461 students to the 47,837 students who had already qualified in the spring. Moreover, the number of students meeting the Benchmarks increased 17% in English, 24% in mathematics, 23% in reading, and 23% in science. For more than 10,000 students who tested in the fall but not the spring, the senior retest opportunity was their only school-day ACT opportunity, and this group was disproportionately Black / African American, Hispanic, and low income.

### *7.5.2 Using ACT Scores as Measures of Prior Achievement for Program Evaluation*

ACT scores may be used as measures of prior achievement to statistically control for differences among program participation groups when evaluating the effectiveness of specific educational programs in observational studies. Two examples are provided in this section.

The first example is a study that Noble and Sawyer (2013) conducted to examine whether taking developmental courses in college benefits students, in the sense that they are more successful than they would have been if they had not taken developmental courses. The study estimated students' chances of success in a variety of college outcomes (including grade in the standard-level course) using ACT test score, enrollment status (full- or part-time), and college type (2- or 4-year) for students who first took the developmental course and then took the standard-level course and for those who enrolled directly in the standard-level course. The chances of success for the students in the two groups were then compared over the range of overlapping ACT scores. Like other studies of this kind, this study indicated that students who took developmental courses were less successful as a group than those who did not take developmental courses with respect to GPA, persistence over time, and degree completion within a fixed time period. However, while students who took developmental courses could take up to a year longer to finish their bachelor's degrees than students who did not take development courses (6 years versus 5), completion rates were generally similar between the two groups. The study also found that certain student subgroups benefited from developmental courses (as indicated by their ACT scores, developmental course grades, and enrollment status) and that part-time students benefited more than full-time students.

The second example comes from two studies undertaken to compare the short- and long-term college outcomes of students who took dual enrollment coursework in high school and those who did not (Crouse & Allen, 2014; Radunzel et al., 2014). The studies revealed that students entering with dual credit hours were generally more academically able (as measured by ACT scores and high school rank) than students who had not taken dual credit in high school. However, these studies were motivated by concerns that dual credit courses might not sufficiently prepare students for college coursework, so the researchers also examined how the two groups differed on other student and school characteristics related to college success. After statistically controlling for ACT scores and other student and school characteristics, the researchers found that students entering college with dual enrollment credit performed as well as those with no dual credit in terms of grades earned in subsequent college courses. Additionally, dual credit students were generally more likely to complete a college degree in a timely manner.

These studies demonstrate the value of including ACT scores as measures of prior achievement in program evaluation studies. For more details on the examples provided, see the full reports.

# Chapter 8
# Growth Models Using ACT Test Scores

Understanding student growth models can help students, parents, educators, and practitioners make better use of ACT® test data. Growth models may be used to answer several important questions:

- How does the growth of students from my school compare to national growth averages?
- How much do my students need to grow to reach their ACT score goals?
- How much do ACT test scores typically increase over a 1-year period?
- Which high school courses have the strongest relationships with student growth?

Growth models that incorporate scores from various ACT assessments (e.g., PreACT® and the ACT) can be used to measure progress—both for individuals and groups of students. Measures of student growth can be used to inform teaching practices and to assess the effectiveness of new programs and interventions. This chapter first distinguishes gain-based models from conditional status models. Subsequent sections discuss resources that are available for implementing growth models based on the ACT test, summarize research explaining variations in student growth, discuss using growth models for evaluating programs and school effectiveness, and summarize research about ACT test–retest statistics. For more information, refer to the ACT Growth Modeling Resources provided online, which include a lookup table of student growth percentiles for different ACT assessments and grade levels.

## 8.1 Distinguishing Gain-Based Models From Conditional Status Models

Methods for describing student- and group-level growth include ones based on gain scores, trajectories, achievement level transitions, residual gains, projections, conditional growth percentiles, and multivariate models. Castellano and Ho (2013) describe each type of growth model and classify these methods by their underlying statistical foundations into one of three categories: gain-based models, conditional status models, and multivariate models. All three categories of growth models can be used to analyze ACT test scores. However, as this chapter describes, the ACT most directly supports gain-based and conditional status models.

Gain-based and conditional status models support contrasting perspectives on growth. Understanding the two models is essential for accurately selecting and using growth models. Gain-based and conditional status models are fundamentally different in two ways: statistical foundation and reliance on common score scales.

### 8.1.1 Statistical Foundation

A gain score is the arithmetic difference between two scores at different time points. Gain-based models express growth as the difference in test scores over time and are meant to answer the question "How much has a student learned on an absolute scale?" (Castellano & Ho, 2013, p. 35). Gain scores may be extrapolated to future time points to support growth predictions. Trajectory models are a type of gain-based model meant to answer the question "If this student continues on this trajectory, what will the student most likely score in the future?"

In contrast, conditional status models address the question "How well did a student score relative to the scores of peers with similar score histories?" While gain-based models attempt to describe growth in an absolute sense, conditional status models attempt to describe growth relative to that of a student's peers. Conditional status models support normative interpretations of student growth.

Conditional status models are often based on regression methods that establish expectations for students' test scores based on their past scores. Comparing actual test scores to expected test scores allows educators to determine whether students have met expectations for "normal" growth. Popular forms of the conditional status model include the student growth percentile (SGP) model and the residual gain model. Similar to gain-based models, conditional status models can describe student growth and predict future test scores. SGP is the model used for the ACT Growth Modeling Resources, as discussed in Section 8.2.

### 8.1.2 Reliance on Common Score Scales

Gain-based models require that test scores from multiple time points share a common scale. This can be achieved through vertical scaling (e.g., test scores from two grade levels placed on the same scale) or by using the same test at multiple time points. When the tests share a common scale, the difference in test scores is meaningful, supporting gain-based models. PreACT and ACT are examples of different tests that are aligned to the same scale.

In contrast, conditional status models do not require that the tests have a common scale. Instead, regression methods are often used to make conditional status models operational. Expectations for test scores ($Y$) are based on a prior test score ($X$) or set of prior test scores. As in regression, $Y$ and $X$ are not required to be on the same scale. Test scores that are on a common scale may be used within both gain-based models and conditional status models. Conditional status models are flexible: A model can use scores from a single prior year or a collection of scores from multiple prior years.

## 8.2 ACT Growth Modeling Resources

To help users implement growth models based on different ACT assessments, ACT provides SGP lookup tables for various assessment combinations and grade levels. ACT Growth Modeling Resources are updated periodically to reflect changes in student growth over time and add more combinations of assessments and grade levels.

The SGP model describes a student's current achievement compared to that of other students with similar prior achievement scores. This model expresses growth as a percentile rank relative to that of the student's academic peers. It is meant to answer the question "What is the percentile rank of a student's current score compared to the scores of students with similar score histories?" For example, a student earning an SGP of 75 performed as well as or better than 75% of the student's academic peers with similar score histories. SGPs provided by ACT are expressed as whole number values from 1 to 99.

Like other conditional status models, the SGP model accommodates multiple prior test scores and does not require test scores from multiple time points to share a common scale. SGPs are often calculated using quantile regression (Koenker, 2005). This method for calculating SGPs does not require linear relationships between prior and current test scores, nor does it require constant variance across prior scores. Open-source software that estimates SGPs using quantile regression methods is available in the SGP package for R (Betebenner et al., 2014).

Many states and school systems use the SGP model to describe student growth, predict future test scores, and examine differences in growth across student groups (e.g., by race/ethnicity, gender, economic status, and disability status). Measures of aggregate growth include the mean and median SGP. Research suggests that mean SGP may have advantages over median SGP in terms of efficiency, greater alignment with expected values, and greater robustness to scale transformations (Castellano & Ho, 2015).

The mean SGP can be used to identify relative growth differences across classrooms, schools, districts, and other groups of interest. When comparing mean SGPs across groups, consider whether differences in the composition of the groups could explain differences in mean SGP. For example, a school serving students from low-income families might be expected to have a lower mean SGP than a school serving students from high-income families.

The ACT Growth Modeling Resources include a lookup table that shows the SGP value (ranging from 1 to 99) associated with each combination of current-year test score and prior-year test score. The lookup table provides an estimate of the SGP for each possible combination of scores for the same sections of the test for various growth periods. As of spring 2026, the SGP lookup table includes the following assessment scenarios with the ACT test:

1. PreACT 8/9 to ACT, Fall Grade 9 to Spring Grade 11
2. PreACT 8/9 to ACT, Spring Grade 9 to Spring Grade 11
3. PreACT to ACT, Fall Grade 9 to Spring Grade 11
4. PreACT to ACT, Spring Grade 9 to Spring Grade 11
5. PreACT to ACT, Fall Grade 10 to Spring Grade 11
6. PreACT to ACT, Spring Grade 10 to Spring Grade 11
7. PreACT Secure to ACT, Spring Grade 9 to Spring Grade 11
8. PreACT Secure to ACT, Spring Grade 10 to Spring Grade 11

SGPs are provided for English, mathematics, reading, science, ELA (English language arts), English + reading, STEM (science, technology, engineering, and mathematics), legacy Composite score (based on English, mathematics, reading, and science), and Composite score (based on English, mathematics, and reading). When interpreting SGPs, always consider the reference group used to estimate the model. The SGPs are based on large samples of students from across the country who tested in recent years through school-day testing programs and who had the test score histories required to compute the SGP. Because these samples do not fully reflect the population of students who participate in ACT school-day testing, the samples are weighted to better match that population with respect to student demographics (race/ethnicity and gender), school poverty level, school type (public or private), and school mean ACT score.

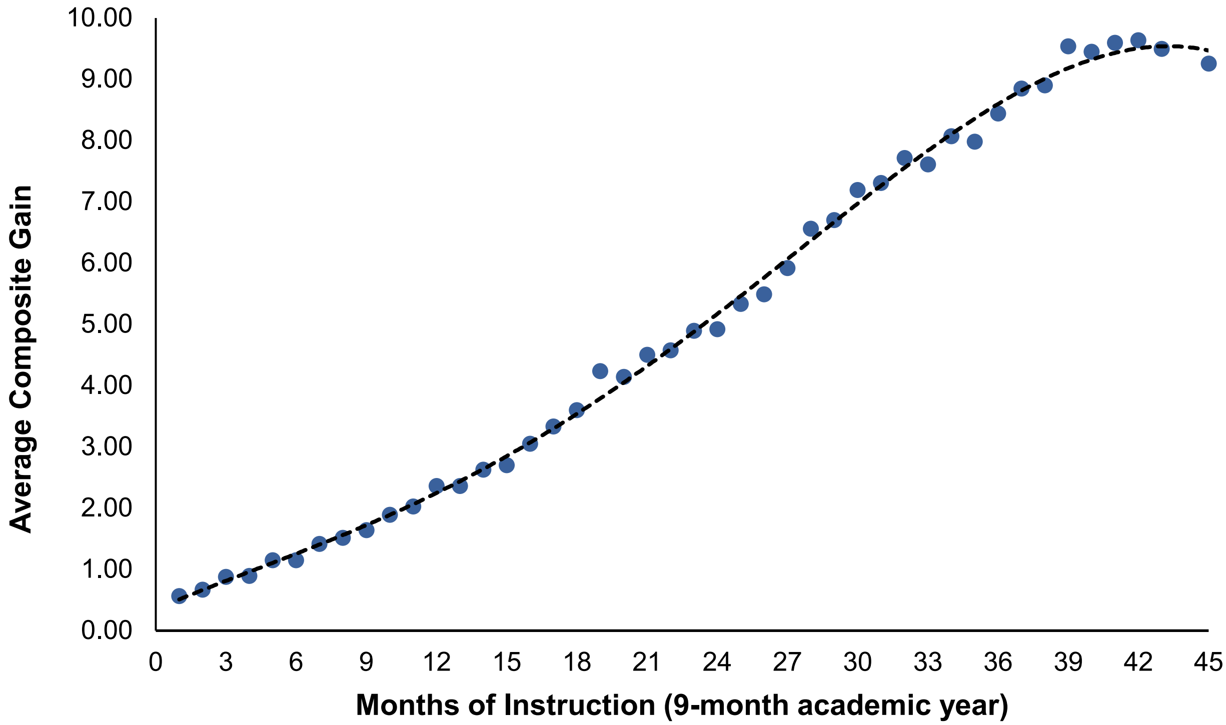
## 8.3 Explaining Variation in Student Growth

Academic growth based on ACT test scores varies across students, schools, and other student groups. Some of the variation in growth can be explained by factors such as instructional time and high school coursework and grades. This section summarizes research that explains some of the variation in student growth.

### *8.3.1 ACT Score Gains by Months of Instruction*

Camara and Allen (2017) used longitudinal data to examine the relationship between instructional time and changes in ACT scores. The sample included over 2.8 million test–retest instances for students from the 2016 ACT-tested graduating class. This research captured typical test–retest periods (e.g., April Grade 11 to October Grade 12) and much longer test–retest periods (e.g., Grade 7 to Grade 12), enabling the researchers to examine ACT score gains across multiple years of instruction. Results indicated that ACT scores steadily increased with more instructional time (Figure 8.1). ACT Composite scores generally increased by 0.20 to 0.25 points per month of instruction, though the increment was larger for shorter periods (1–3 months), perhaps due to practice effects. Over a 4-year period (36 months of instruction), students gained about 8.5 ACT Composite score points on average.

**Figure 8.1.** Average Gain in ACT Composite Score by Months of Instruction




*Note.* Months of instruction are based on time between ACT tests, not counting June, July, and August.

Allen, Mattern, and Camara (2020) used data from the 2017–2018 and 2018–2019 ACT testing years to estimate per-month gains in ACT scores. The analysis was based on 482,413 students who first took the ACT in Grade 11. The researchers used linear regression to estimate the gain in ACT scores (from first to second test) associated with each additional month that school was in session and also the gain (or loss) in ACT scores associated with each additional month that school was not in session. Although positive gains were observed while school was in session, the researchers observed losses while school was not in session, providing evidence of summer-learning loss. The average per-month gains (or losses) varied by test section:

- English: 0.310 points while school in session, −0.159 while not in session
- Mathematics: 0.193 points while school in session, −0.221 while not in session
- Reading: 0.182 points while school in session, −0.023 while not in session
- Science: 0.186 points while school in session, −0.164 while not in session

### *8.3.2 Predictors of Long-Term Growth*

#### Estimated Benefits of Additional High School Coursework and Improved Course Performance in Preparing Students for College

Strategies for increasing academic growth and improving college readiness include taking more rigorous college-preparatory courses and expending more effort in these courses. Sawyer

(2008) examined how taking additional courses and earning higher grades were associated with high school students' academic preparation for college, using data from students who took ACT Explore in Grade 8, ACT Plan in Grade 10, and the ACT in Grades 11 or 12.

Students' background characteristics, ACT Explore scores, high school attended, high school coursework, and high school grades were all related to ACT scores, but ACT Explore scores were by far the most strongly related. Taking more standard or advanced courses in high school and earning higher grades were associated with higher ACT scores even for the students who had high ACT Explore scores. There was significant variation in average ACT scores across high schools, even after accounting for differences in their students' characteristics. The apparent benefit of additional standard coursework, advanced/honors coursework, and higher grades also varied significantly among high schools.

Bassiri (2014) replicated Sawyer's (2008) study using a more recent cohort of students (high school graduates of 2013 who took ACT Explore in Grade 8) and using updated values for the College Readiness Benchmarks. The source data contained records for 399,642 students from 6,228 high schools. In contrast to Sawyer's study, accelerated, honors, or advanced courses were excluded from the predictive models in Bassiri's study because there were large percentages of missing data (39%–56%). In general, Bassiri's findings were consistent with those of Sawyer's study. However, there were a few exceptions, such as English coursework was a significant predictor of ACT English scores, and social studies coursework was a significant predictor of ACT reading scores.

### Predictors of Academic Growth in Secondary School Among Academically Advanced Youth

Academically advanced youth took the ACT test in Grade 7 for academic talent searches and again in Grades 11 or 12 for college admissions (Allen, 2016a), enabling an investigation of predictors of growth during secondary school. Wai and Allen (2019) tested whether variation in academic growth among academically advanced youth is explained by socio-demographics, high school characteristics, coursework taken, high school GPA, Holland-type vocational interests (Holland, 1997), and extracurricular activities.

The study examined predictors of student growth over a long period: from Grade 7 to Grades 11 or 12. Predictors of growth included malleable factors such as high school coursework and grades as well as background variables such as race/ethnicity and family income. All variables combined—socio-demographics, interests, high school characteristics, high school coursework and GPA, and extracurricular activities—explained 25% of the variance in academic growth. Variation in growth was observed across racial/ethnic, gender, and family income groups. Students attending Catholic and private schools had the highest growth, whereas homeschooled students and students attending low-income public schools exhibited lower growth. Malleable factors associated with higher growth included earning higher grades in high school courses, taking elective high school courses in STEM areas, and taking advanced, accelerated, or honors courses. Students with high levels of Investigative and Conventional interests on the Holland scale (1997) had higher growth.

### 8.3.3 Group Differences in Growth

#### Academic Growth Patterns for English Learners and Students With Disabilities

Bassiri and Allen (2012) examined differences in growth for English learners (ELs) and students with disabilities (SWD). This study used longitudinal data for 103,725 students who took ACT Explore in Grade 8, ACT Plan in Grade 10, and the ACT in Grades 11 or 12. Across the test sections, the average ACT Explore, ACT Plan, and ACT scores were highest for the reference group and lowest for the EL and students enrolled in special education courses. Among SWD, the average scores across test sections were lower for students with cognitive or learning disabilities than for students with physical disabilities. In some cases, the growth measures revealed a different pattern. For example, compared to the reference group, ELs had consistently higher growth in mathematics and science between Grades 8 and 10; SWD exhibited above-average growth in reading between Grades 10 and 11/12; and students with cognitive or learning disabilities exhibited above-average growth in reading and science between Grades 10 and 11/12.

Most of the growth differences, while statistically significant, were small in magnitude. For example, students with a cognitive or learning disability increased their scores 0.28 points less than the average increase in scores in mathematics between Grades 8 and 11/12, in which the average gain for the reference group was 5.1 points (16.5 on Explore to 21.6 on the ACT). The 0.28 growth difference was only about 5% of the reference group's overall gain and so is small in magnitude. While many of the growth differences were not statistically significant or were statistically significant but small in magnitude, some differences were more striking: Compared to their reference group peers, ELs increased their scores an average of 0.81 points less than the average gain in English between Grades 10 and 11/12 but grew nearly 0.75 points more than the average gain in mathematics between Grades 8 and 11/12. Students in special education courses grew nearly a full point less than the average in English and students in the "other disability" group grew more than half a point less than the average in mathematics between Grades 8 and 11/12.

#### Academic Growth Patterns of First-Generation College Students

Many college students are first-generation students, meaning that neither parent attended college. First-generation college students tend to have lower scores on college readiness assessments and be less successful in completing their postsecondary programs than students whose parents went to college. Bassiri (2016a) investigated the extent to which gaps in their test scores might begin in middle school. This study used longitudinal data for approximately 282,000 students who took ACT Explore in Grade 8, ACT Plan in Grade 10, and the ACT in Grades 11 or 12. Four groups of students were identified according to their parents' highest grade level: no college experience (first-generation), some college experience, bachelor's degree, or graduate degree. Following the same methodology as Bassiri and Allen (2012), ACT researchers averaged residual gain scores for these same four groups to form measures of aggregate growth for each of these groups and each grade-level period.

Across the test sections, the average ACT Explore, ACT Plan, and ACT scores were positively associated with parents' highest grade level. Both first-generation students and students whose parents had some college experience had statistically significant ($p < .01$) and negative mean residual scores in all test sections across all grade spans. Students whose parents had at least a bachelor's degree had statistically significant positive mean residual scores. Across test sections, it appeared that growth differences by parents' highest grade level became more pronounced over time. That is, the mean residuals were larger in later grades, indicating that differences across socioeconomic levels increased over time.

## 8.4 Using Growth Models for Evaluating Programs and School Effectiveness

ACT-based growth models can support analyses of program evaluation and school effectiveness by comparing students' observed achievement with expected achievement based on prior test performance. For both uses, it is important for students with ACT scores to closely represent the relevant student population, such as in school-day testing contexts. Aggregated growth results, such as mean SGPs or related value-added indicators, can help schools, districts, and programs examine whether students or groups are growing more or less than their academically similar peers.

### *8.4.1 Program Evaluation*

ACT scores and growth measures can be used as one source of evidence for evaluating instructional programs, interventions, or policies when the ACT content aligns with the intended outcomes of the program. While ACT-based growth models can strengthen these evaluations by incorporating prior achievement, the results should be interpreted alongside other information routinely collected by schools or programs. For broad validity considerations related to using ACT scores in program evaluation, see Chapter 7.

### *8.4.2 Measures of High School Effectiveness*

ACT-based growth models and value-added models have both been used to examine school effectiveness by estimating whether students' observed achievement is higher or lower than would be expected according to the students' prior achievement levels. Previous ACT research has shown that different accountability and value-added models can lead to different conclusions about a school's effectiveness (Allen et al., 2009) and that, in most cases, estimated school effects do not differ significantly from the average school effect, underscoring the importance of reporting the statistical uncertainty around these estimates. Subsequent research found that value-added scores based on longer growth periods (e.g., Grades 8–12) were better able to distinguish school effects, that prior academic achievement was positively related to value-added measures, and that school poverty (proportion of students eligible for free or reduced-price lunch) and proportion of racial/ethnic minority students were negatively related, whereas class size and proportion of students tested had weaker associations (Bassiri, 2015).

A related study found that high school value-added measures based on ACT scores had small but statistically significant relationships with college enrollment, retention, and first-year course grades in selected core content areas, after accounting for student- and school-level characteristics (Bassiri, 2016b). Most of the variance in these college outcomes was attributable to student characteristics rather than to high schools or colleges, as reflected in small but statistically significant intraclass correlation coefficients. Taken together, these findings suggest that ACT-based growth measures can provide useful evidence for school improvement discussions, but these results should be interpreted with attention to model choice, growth period, prior achievement measures, student composition, representativeness, and statistical uncertainty.

## 8.5 ACT Retest Statistics

Many students take the ACT more than once, and ACT retest statistics are useful for understanding how much gain that students typically achieve. The list at the end of this section summarizes the research on ACT retest statistics for the legacy ACT test. While ACT retest statistics for the enhanced ACT are expected to be similar to those for the legacy ACT, future studies will examine retest statistics specifically for the enhanced ACT.

Harmston and Crouse (2016) examined the trends associated with students who retest, focusing on the number of times students took the ACT test and the time elapsed between their tests. Score gains were highest for students who initially had low scores and for students who first tested in Grade 10. Overall, ACT Composite score gains were modest. The researchers concluded that students should consider retesting if they believe their test scores do not accurately reflect their skills and knowledge.

Gains from the first to second ACT test have also been examined for over 772,000 students from the ACT-tested graduating class of 2013 who took the ACT two or more times (Camara & Allen, 2017). The results showed that 57% of students improved their ACT Composite score, 21% saw no change, and 22% saw a decrease in their ACT Composite score. For students with an initial ACT Composite score between 13 and 29, the typical gain in ACT Composite score from the first to second test was 1 point.

In a study of students from the ACT-tested graduating class of 2018 who retested, Harmston (2020) modeled students' chances of no ACT Composite score gain (including score drops) and gains of 1, 2, and 3 or more points upon their second testing attempt as a function of student educational performance and behavioral attributes. The variables that were identified as having the strongest relationships with score gains included initial ACT Composite score, grade level at the time of first testing, time between the two testing events, the interaction between initial ACT Composite score and time between tests, high school GPA, plans to take physics in high school, plans to take calculus in high school, and plans to take one or more accelerated, honors, or advanced courses in high school. Results from this study were used to develop an ACT web application that enables users to enter student-specific criteria to calculate the likelihood of increasing a student's Composite score.

Another study examined score gains for students in the ACT-tested graduating class of 2022 who retook the ACT (Allen, 2022). About 35% of the students in this graduating cohort tested at least twice: 21% tested twice, 8% three times, 4% four times, and 3% five or more times. Findings include the following:

- The average ACT Composite score gain from first to last test was 1.14 points. The average gain varied across test sections, from a low of 0.86 for mathematics to a high of 1.66 for English.
- Score gains increased with more time between tests and were greater when the time between tests did not include a summer break. For example, the average ACT Composite gain was 0.6 when the tests were taken 6 months apart and included a summer break (e.g., April to October), but the average gain was 1.1 when the tests were taken 6 months apart and did not include a summer break (e.g., December to June).
- Score gains were higher for students who took challenging high school courses and earned good grades.
- Score gains were compared by family income level and racial/ethnic group. The differences in score gains across the demographic groups were small but consistent with long-standing achievement gaps.
- Students who retested more than once usually saw additional increases in their ACT scores, but the score increases shrunk with each additional test. For example, for students who tested four times, the average Composite score increased by 1.05 points from the first to the second test, 0.59 points from the second to the third test, and 0.32 points from the third to the fourth test.

# Chapter 9
# Other ACT Components

## 9.1 The ACT Interest Inventory

The primary purpose of the ACT® Interest Inventory is to stimulate and facilitate exploration of personally relevant educational and occupational options. Given the important decisions and choices students must make as they navigate the transition from high school to college or career, exploring the self in relation to education and career options is critical. Using their Interest Inventory results, students can explore programs of study and occupations in line with their interests.

The ACT Interest Inventory consists of 72 items and provides scores on six scales that parallel Holland's (1997) six types of interests and occupations (see also Holland et al., 1969). Scale names (and parallel Holland types) are Science & Technology (Investigative), Arts (Artistic), Social Service (Social), Administration & Sales (Enterprising), Business Operations (Conventional), and Technical (Realistic). Each scale consists of common, everyday activities that are both familiar to students and relevant to work (e.g., study biology, help settle an argument between friends, sketch and draw pictures). The activities have been carefully chosen to assess basic work-relevant interests while minimizing the effects of sex-role connotations. Because males and females obtain similar distributions of scores, combined-sex norms are used to obtain sex-balanced scores. Readers seeking additional information about the ACT Interest Inventory are encouraged to consult the ACT Interest Inventory Technical Manual (ACT, 2023).

### *9.1.1 Reporting Procedures*

#### Student Report

To facilitate educational and occupational exploration, results reported to students are expressed visually in work-world terms. Extensive research (much of it cited in Prediger, 1996) indicates that two orthogonal work-task dimensions (Data/Ideas and People/Things) underlie Holland's hexagonal model of interests and occupations (Holland, 1997; Holland et al., 1969). Thus, a two-dimensional space can display both a comprehensive set of occupations and the results of measured interests.

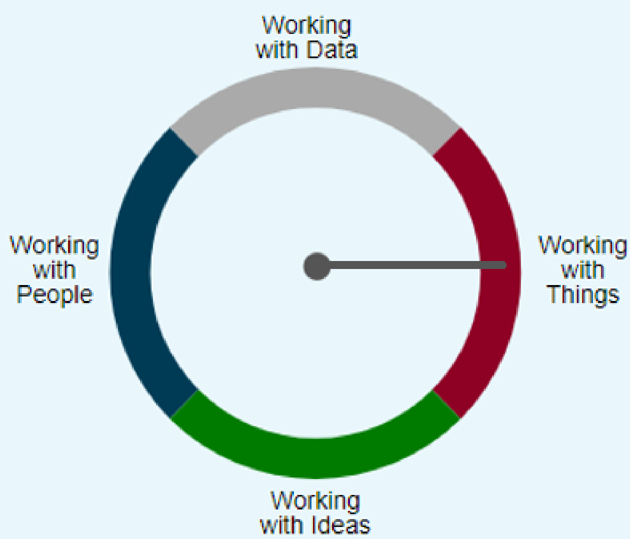
The ACT Interest Inventory reports results to students in two ways. First, it displays the results on the Career Connector. Second, it includes a short list of occupations that primarily involve the kinds of basic work tasks that the student prefers. The Career Connector is a two-dimensional figure with four compass points labeled Working with People, Working with Data, Working with Things, and Working with Ideas (Figure 9.1; see ACT, 2023, for definitions). The Career Connector summarizes the pattern of scores on the six ACT Interest Inventory scales and visually displays it by pointing at one or two compass points, which indicates the direction of a student's interests. For example, the point on the Career Connector that falls between ideas and people suggests that the student primarily enjoys activities involving these kinds of work tasks. The Career Connector is derived from the ACT Career Map, an empirically based system for summarizing basic similarities and differences between groups of occupations with respect to their relative involvement with people, data, things, and ideas. As described below, the Career Map links people to occupations by providing a visual display of personalized and actionable assessment results.

**Figure 9.1.** Example ACT Interest Inventory Results From the Student Report on MyACT

## College and Career Planning

### Interest-Work Preferences

Occupations differ widely in how much they involve working with four basic work tasks: Data, Ideas, People and Things. The results of your Interest Inventory point to the kinds of basic work tasks you prefer.




According to your results, you enjoy working with things. Here are a few examples of occupations involving this kind of work:

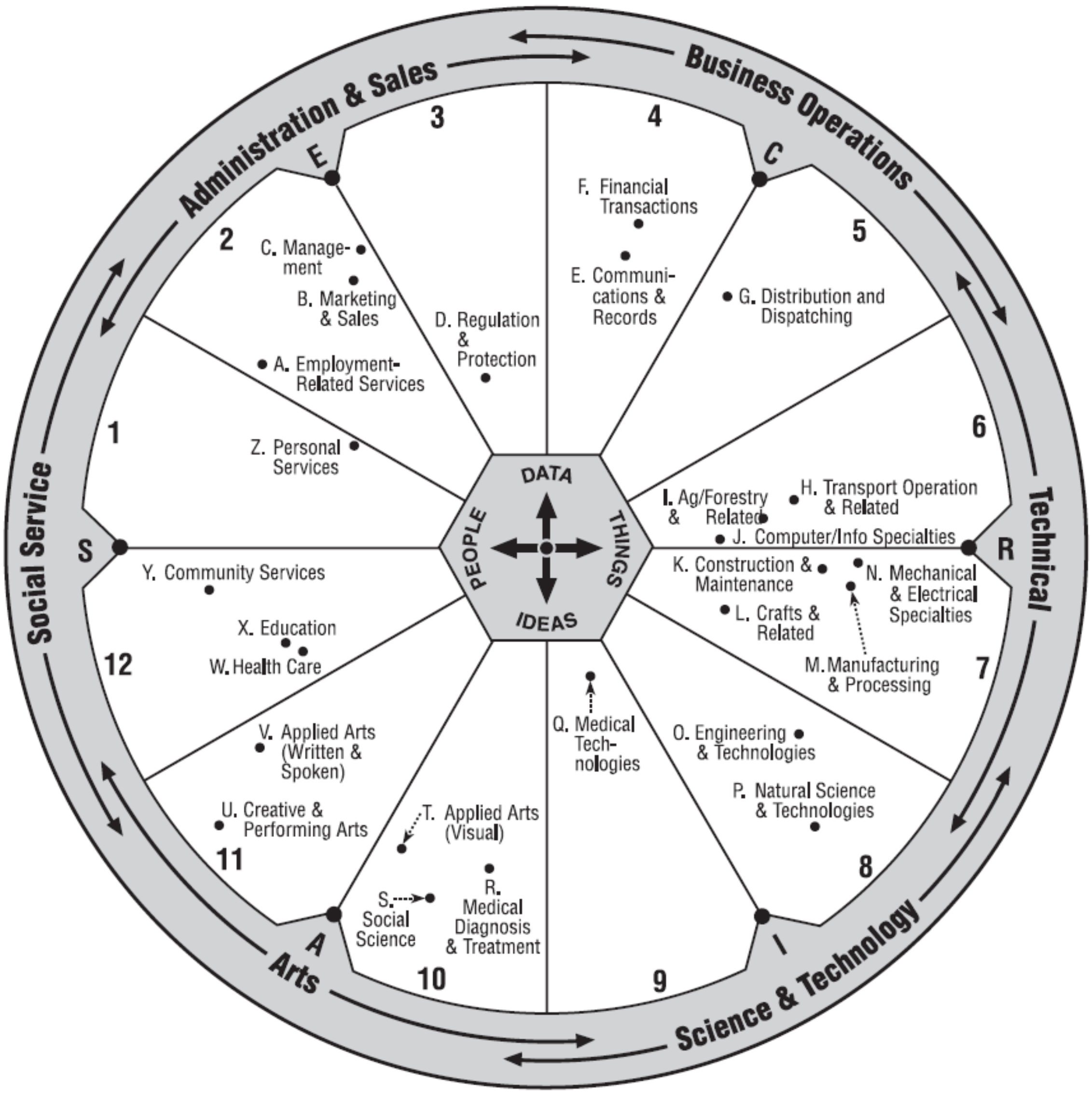
- Air Traffic Controller
- Broadcast Technician
- Computer Programmer
- Forester
- Machinist/Tool Programmer

For more details on how your interests align with your major choice and future career check out the plans tab.

### Career Map

The ACT Career Map (Figure 9.2) provides a simple yet comprehensive overview of the world of work and a visual means of linking ACT Interest Inventory scores to career options. The 26 career areas (groups of related occupations) are located in 12 map regions. Career areas are located on the Career Map according to the relative standing of their member occupations on the Data/Ideas and People/Things work task dimensions. Career area locations are based on extensive and diverse occupational data involving expert ratings, job analyses, and measured interests (ACT, 2023; Prediger & Swaney, 2004). The purpose of the work and the work setting were also considered when the career areas were developed.

**Figure 9.2.** The ACT Career Map




Although ACT took care to make each career area as homogeneous as possible, there is scatter across the occupations in each area. The scatter could be reduced by using more areas, but the Career Map was constructed for applied purposes and is not meant to provide a precise scientific statement. As can be seen in Figure 9.2, career area locations generally make good theoretical and common sense.

A student's pattern of ACT Interest Inventory scores is converted to map regions, and the career areas that align with the student's score pattern are reported, allowing for focused exploration of occupations that fit the student's interests. The method for converting scores to map regions is summarized in Appendix C of the *ACT Interest Inventory Technical Manual* (ACT, 2023).

### *9.1.2 Psychometrics*

The *ACT Interest Inventory Technical Manual* (ACT, 2023), which presents a wide range of information about the inventory, includes the following topics:

- inventory items, scales, and interpretive aids
- development of items and norms
- reliability (internal consistency and test-retest stability)
- validity (convergent and discriminant evidence, item and scale structure, interest-environment fit, and success outcomes)

Internal consistency reliability coefficients for the six 12-item scales based on a Grade 12 sample ($N$ = 20,000) ranged from .84 to .91 (median = .87). Validity evidence is extensive, including discriminant validity evidence based on the score profiles of 648 career groups (representing over 79,000 college major and occupation incumbents) and scale-structure evidence based on multiple samples ($N$ = 60,000).

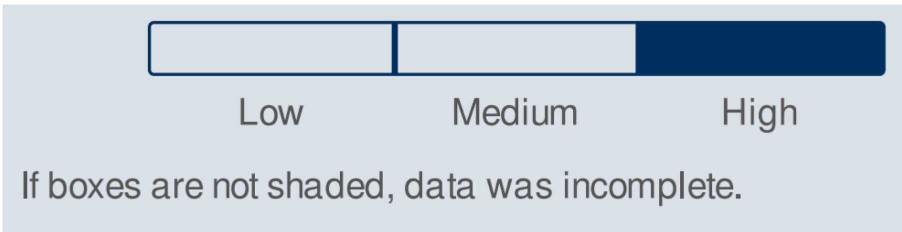
### *9.1.3 Interest-Major Fit*

Interest-major fit is derived from two data elements collected during ACT test registration: the student's ACT Interest Inventory scores and the major the student plans to pursue. Interest-major fit measures the strength of the relationship between the student's profile of ACT Interest Inventory scores and the profile of interests of students in the student's planned major. Interest profiles for each of the 294 majors on the ACT registration list are based on a large national sample of undergraduate students with a declared major and a GPA of at least 2.0. A student's major was determined in the third year for students in 4-year colleges and in the second year for students in 2-year colleges.

Interest-major fit scores range from 0 to 99. The higher the score, the better the fit. Using data from a large national sample, ACT established three levels of fit based on the empirical relationships between the interest-major fit scores and the proportion of students who persisted in their college major. In the college score report, shading along an interest-major fit bar indicates the strength of fit: low, medium, or high (Figure 9.3).

Empirical evidence indicates that the fit between students' interests and their college majors is important to understanding and predicting student outcomes. Research involving the ACT Interest Inventory suggests that if students' measured interests (i.e., patterns of interest scores) are similar to the interests of people in their chosen college majors, they will be more likely to persist in college (Allen & Robbins, 2008; Tracey & Robbins, 2006), remain in their majors (Allen & Robbins, 2008), and complete their college degrees in a timely manner (Allen & Robbins, 2010). Even before students declare a major in college, fit between their interests and planned major is a good predictor of whether they will follow through on their college major plans (ACT, 2013b). The value of interest-major fit is not limited to the ACT Interest Inventory or to the outcomes listed above. A large-scale meta-analysis, which analyzed data gathered over a 60-year period and included a range of outcome and interest measures (including the ACT Interest Inventory), found that interest-environment fit is related to persistence and performance in both academic and work settings (Nye et al., 2012). Additional information on research involving the ACT Interest Inventory and interest-major fit is described in ACT (2023).

**Figure 9.3.** Example Interest-Major Fit Bar From the ACT College Report




## 9.2 The MyACT Profile

Other student information is collected as part of ACT registration to broaden the information base for both students and colleges. The MyACT Profile collects information and makes it available in a systematic form prior to college enrollment. The development of the MyACT Profile was influenced by the educational context in which it evolved. The information made available to students helps inform their educational choices, and the information made available to colleges provides a more comprehensive picture of the quality of education students received in high school.

The MyACT Profile contains several sections, each of which is discussed below. The items were developed by ACT staff with input from personnel representing a variety of postsecondary educational institutions. Items are revised from time to time to meet institutions' changing needs for different types of data. Current MyACT Profile questions and response options are available on the ACT website through the MyACT Profile page.

### *9.2.1 Demographic Information and ACT Recruit Me*

In this section of the MyACT Profile, students have the option of providing common demographic information like age, race/ethnicity, gender, language, and parent/guardian education level. ACT collects and reports this information in accordance with reporting guidelines issued by the U.S. Department of Education, which do not include collecting ethnic background information at a more detailed level. This information is released to the colleges that receive student scores only if students request it. This section also provides students with the option to opt in to ACT Recruit Me. This is a voluntary program that helps students connect with colleges and scholarship agencies seeking prospective students and makes them aware of other educational opportunities. Participating organizations have agreed to use student information only to share information about their programs. ACT research shows that participation in ACT Recruit Me benefits students by giving them greater access to information about colleges and programs of study, thereby expanding their potential for future success.

### 9.2.2 Your High School Résumé

This section requests information about students' high school performance (overall GPA, class rank), their program of high school courses, the number of students in their graduating class, and the type of high school they attended (public/private). If a student will graduate from a home school, this section asks how many years of high school homeschooling the student will have completed by graduation.

### 9.2.3 Your High School Courses & Grades

Most colleges, universities, and state agencies seek or require information from applicants on performance in a wide range of high school courses. To meet this need, ACT—in consultation with a representative group of personnel from postsecondary educational institutions—developed a list of 30 high school courses. In this section of the MyACT Profile, students are asked to report the grades they earned in these 30 courses, spanning six academic areas: English, mathematics, natural sciences, social studies, languages, and arts. ACT calculates a grade point average (GPA) on an unweighted 4.0 scale based on a student's responses. High school GPAs based on self-reported grades are shown on the ACT College Report for English, mathematics, natural sciences, and social studies. For each subject area, students are asked the number of years of study expected by graduation. Students are also asked whether they have been enrolled in advanced placement, accelerated, or honors courses in any of five areas (English, mathematics, social studies, natural sciences, and foreign languages). Because high school grades depend on both academic aptitude and personal characteristics such as persistence and study habits, these self-reports provide useful estimates of future academic achievement. Validity evidence for self-reported high school grades is discussed in Chapter 7.

### 9.2.4 College Plans

Questions about students' college plans are essential to the planning done by colleges. In this section, questions first ask about students' plans for financing a college education. Students are asked to estimate their families' annual income and indicate whether they intend to apply for financial aid and/or to work while in college. Student responses to these questions can be useful to college financial aid officers.

Information about how students choose a college can be useful to college personnel. This section also asks students about their enrollment plans (full-time or part-time); their preferred institution type (public/private, 2-year/4-year), size, and location; and their preferred living accommodations. The student is asked to rank institutional preferences in order of importance.

With each new entering class, colleges must be prepared to provide individualized assistance to support the academic development of their students. Additional questions in this section provide a list of specific academic skills and ask students what areas they need assistance in; areas include educational and occupational planning, writing, reading, study skills, and mathematics. By providing such information, students alert the college to their individual needs. This section also includes questions about student interest in college programs designed for enriched or accelerated academic work.

Entry into postsecondary education, as well as progress through such education, requires that students make important decisions and choices. Even tentative choices are important in that they provide a foundation for (and often place limits on) future finalized choices. The MyACT Profile questions provide opportunities for students to indicate information such as their intended college major, current occupational choice, and level of educational aspirations. This information is useful in evaluating how realistic a student’s choices are, as well as providing colleges with data that can be used for planning educational programs that meet the needs of all their students.

### *9.2.5 Extracurriculars*

In this section, students are asked to indicate high school activities in which they have participated and activities in which they plan to participate during college. Students select from a list of activities such as athletics, drama, music, student government, student publications, and special-interest clubs. Information about students’ current activities can be useful for connecting students to specific college opportunities. Information about the prospective plans of their incoming students is valuable to colleges seeking to develop appropriate extracurricular programs. From the student perspective, presenting their extracurricular activities and plans is another way of communicating their unique patterns of interests and skills.

# References

ACT. (1989). *Using concordance tables.*

ACT. (2005). *Courses count: Preparing students for postsecondary success* [Policy report]. https://www.act.org/content/dam/act/unsecured/documents/CoursesCount.pdf

ACT. (2007). *ACT National Curriculum Survey 2007.*

ACT. (2009). *ACT National Curriculum Survey 2009.* https://www.act.org/content/dam/act/unsecured/documents/NationalCurriculumSurvey2009.pdf

ACT. (2010). *Mind the gaps: How college readiness narrows achievement gaps in college success*. https://www.act.org/content/dam/act/unsecured/documents/MindTheGapsSummary.pdf

ACT. (2013a). *ACT National Curriculum Survey 2013. https://www.act.org/content/dam/act/unsecured/documents/Natl-Scores-2013-National2013.pdf*

ACT. (2013b). *College choice report part 3: Persistence and transfer*. https://www.act.org/content/dam/act/unsecured/documents/CollegeChoiceReport2012-3.pdf

ACT. (2015). *ACT CAAP guide to successful general education outcomes assessment*.

ACT. (2016a). *ACT National Curriculum Survey 2016.* https://www.act.org/content/dam/act/unsecured/documents/NCS_Report_Web_2016.pdf

ACT. (2016b). *ACT national profile report 2016*. https://www.act.org/content/dam/act/unsecured/documents/P_99_999999_N_S_N00_ACT-GCPR_National.pdf

ACT. (2020). *ACT National Curriculum Survey 2020.* https://www.act.org/content/dam/act/unsecured/documents/NCS_Report_Web_2020.pdf

ACT. (2023). *The ACT Interest Inventory technical manual*. https://www.act.org/content/dam/act/unsecured/documents/ACT-Interest-Inventory-Technical-Manual.pdf

ACT. (2025a). *PreACT 8/9 technical bulletin*. https://www.act.org/content/dam/act/unsecured/documents/PreACT-8-9-Technical-Bulletin-2025.pdf

ACT. (2025b). *PreACT technical manual*. https://www.act.org/content/dam/act/unsecured/documents/PreACT-Technical-Manual.pdf

Allen, J. (2013). *Updating the ACT College Readiness Benchmarks* (Research Report 2013-6). ACT. https://www.act.org/content/dam/act/unsecured/documents/ACT_RR2013-6.pdf

Allen, J. (2015). *Influence of achievement in core high school courses on ACT scores* (Technical Brief 4862). ACT. https://www.act.org/content/dam/act/unsecured/documents/2015-Tech-Brief-Influence-of-Achievement.pdf

Allen, J. (2016a). *How do 7th graders who take the ACT test score as 11th or 12th graders?* [Technical brief]. ACT. https://www.act.org/content/dam/act/unsecured/documents/How-Do-7th-Graders-Who-Take-the-ACT-Score-as-11th-Graders.pdf

Allen, J. (2016b). *Subgroup representation and validity in ACT College Readiness Benchmark development samples* [Technical brief]. ACT. https://www.act.org/content/dam/act/unsecured/documents/5483-tech-brief-subgroup-representation.pdf

Allen, J. (2022). *Six things you should know about ACT score gains from retesting* (Issue Brief R2274). ACT. https://www.act.org/content/dam/act/unsecured/documents/2022/R2274-Six-Things-About-Score-Gains-From-Retesting-10-2022.pdf

Allen, J. (2025). *ACT score norms for students in Grades 7–9* (Technical Brief R2536). ACT. https://www.act.org/content/dam/act/unsecured/documents/R2536-ACT-Score-Norms-for-Students-in-Grades-7-through-9-2025-11.pdf

Allen, J. (2026). *PreACT-tested students in the 2025 ACT-tested graduating class: A summary of performance and growth* (Issue Brief R2605). ACT. https://www.act.org/content/dam/act/unsecured/documents/R2605-PreACT-Tested-Students-in-the-2025-ACT-Tested-Graduating-Class-2026-02.pdf

Allen, J., Bassiri, D., & Noble, J. (2009). *Statistical properties of accountability measures based on ACT's educational planning and assessment* (Research Report 2009-1). ACT. https://www.act.org/content/dam/act/unsecured/documents/ACT_RR2009-1.pdf

Allen, J., & Cruce, T. (2025). *Initial evidence supporting interpretations of scores from the enhanced ACT test* (Research Report R2425). ACT. https://www.act.org/content/dam/act/unsecured/documents/R2425-Initial-Evidence-Supporting-Interpretations-of-Scores-from-the-Enhanced-ACT-Test-2025-06.pdf

Allen, J., Cruce, T., & Dingler, C. (2023). *School-day administration of the ACT test: Removing barriers and opening doors for all students* (Issue Brief R2310). ACT. https://www.act.org/content/dam/act/unsecured/documents/R2310-School-Day-ACT-Administration-Removing-Barriers-for-All-Students-06-2023.pdf

Allen, J., & Dingler, C. (2023). *Examining benefits of school-day ACT senior retesting* (Issue Brief R2289). ACT. https://www.act.org/content/dam/act/unsecured/documents/R2289-Examining-Benefits-of-School-Day-ACT-Senior-Retesting-11-2023.pdf

Allen, J., Mattern, K., & Camara, W. (2020). *Predicting the impact of COVID-19 school closures on ACT test scores: Methods and considerations for states and districts* (Issue Brief R1828). ACT. https://www.act.org/content/dam/act/unsecured/documents/R1828-covid-impact-2020-05.pdf

Allen, J., Mattern, K., & Ndum, E. (2019). An empirically derived index of high school academic rigor. *Educational Measurement: Issues and Practice*, 38(1), 6–15. https://doi.org/10.1111/emip.12236

Allen, J., & Radunzel, J. (2016). *Comparing high school grade point average (GPA) to first-year college GPA* (Data Byte 2016-10). ACT. https://www.act.org/content/dam/act/unsecured/documents/5840-Data-Byte-2016-10-Comparing-High-School-GPA-to-First-Year-College-GPA.pdf

Allen, J., & Robbins, S. (2008). Prediction of college major persistence based on vocational interests, academic preparation, and first-year academic performance. *Research in Higher Education, 49*(1), 62–79. https://doi.org/10.1007/s11162-007-9064-5

Allen, J., & Robbins, S. (2010). Effects of interest-major congruence, motivation, and academic performance on timely degree attainment. *Journal of Counseling Psychology, 57*(1), 23–35. https://doi.org/10.1037/a0017267

Allen, J., Thomas, J., Dreyer, S., Johanningmeier, S., Murano, D., Cruce, T., Li, X., & Sanchez, E. (2026). *Design framework for the ACT enhancements* (Research Report R2519).

ACT. https://www.act.org/content/dam/act/unsecured/documents/R2519-Design-Framework-for-the-ACT-Enhancements-2026-02.pdf

Allen, J. D. (2005). Grades as valid measures of academic achievement of classroom learning. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, *78*(5), 218–223. https://doi.org/10.3200/TCHS.78.5.218-223

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association. https://www.testingstandards.net/open-access-files.html

Arthur, A., Kapoor, S., & Steedle, J. (2020). *Paper and online testing mode comparability: A review of research from 2010–2020* (Technical Brief R1842). ACT. https://www.act.org/content/dam/act/unsecured/documents/R1842-paper-online-testing-modes-2020-12.pdf

Arthur, A., & Thomas, J. (2023, April). *Using eyetracking to address reading rates and speededness on high stakes assessments* [Paper presentation]. Annual Meeting of the National Conference on Measurement in Education, Chicago, IL, United States. https://doi.org/10.13140/RG.2.2.29263.19365

Bassiri, D. (2014). *The forgotten middle: Ensuring that all students are on target for college and career readiness before high school*. ACT. https://www.act.org/content/dam/act/unsecured/documents/ForgottenMiddleSummary.pdf

Bassiri, D. (2015). *Statistical properties of school value-added scores based on assessments of college readiness* (Research Report 2015-5). ACT. https://files.eric.ed.gov/fulltext/ED563776.pdf

Bassiri, D. (2016a). *Academic growth patterns of first-generation college students in grades 8 to 12* (Research Report 2016-6). ACT. https://www.act.org/content/dam/act/unsecured/documents/5881-research-report-2016-6-academic-growth-patterns.pdf

Bassiri, D. (2016b). *Are value-added measures of high school effectiveness related to students' enrollment and success in college?* (Research Report 2016-10). ACT. https://files.eric.ed.gov/fulltext/ED573780.pdf

Bassiri, D., & Allen, J. (2012). *Grade 8 to 12 academic growth patterns for English language learners and students with disabilities* (Research Report 2012-1). ACT. https://www.act.org/content/dam/act/unsecured/documents/ACT_RR2012-1.pdf

Beatty, A. S., Walmsley, P. T., Sackett, P. R., Kuncel, N. R., & Koch, A. J. (2015). The reliability of college grades. *Educational Measurement: Issues and Practice*, *34*(4), 31–40. https://doi.org/10.1111/emip.12096

Betebenner, D. W., Van Iwaarden, A., Domingue, B., and Shang, Y. (2014). *SGP: An R package for the calculation and visualization of student growth percentiles & percentile growth trajectories* (Version 1.2-0.0) [Computer software]. https://cran.r-project.org/web/packages/SGP/index.html

Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In F. M. Lord & M. R. Novick (Eds.), *Statistical theories of mental test scores* (pp. 397–479). Addison-Wesley.

Brookhart, S. M. (1993). Teachers' grading practices: Meaning and values. *Journal of Educational Measurement*, *30*(2), 123–142. https://doi.org/10.1111/j.1745-3984.1993.tb01070.x

Brookhart, S. M. (2015). Graded achievement, tested achievement, and validity. *Educational Assessment, 20*(4), 268–296. https://doi.org/10.1080/10627197.2015.1093928

Camara, W., & Allen, J. (2017). *Does testing date impact student scores on the ACT?* (Technical Brief R1643). ACT. https://www.act.org/content/dam/act/unsecured/documents/R1643-test-date-impact-on-act-scores-2017-06.pdf

Casillas, A. (2013, October). *The other side of college readiness: Behaviors drive academic success* [Paper presentation]. Annual Conference of the National Scholarship Providers Association, Portland, OR, United States.

Casillas, A., Way, J. D., & Burrus, J. (2015). Behavioral skills. In W. J. Camara, R. O'Connor, & K. Mattern (Eds.), *Beyond academics: A holistic framework for enhancing education and workplaces success* (pp. 25–38). ACT. https://www.act.org/content/dam/act/unsecured/documents/RR2015-4-beyond-academics-a-holistic-framework-for-enhancing-education-and-workplace-success.pdf

Castellano, K. E., & Ho, A. D. (2013). *A practitioner's guide to growth models.* CCSSO. https://andrewho.scholars.harvard.edu/sites/g/files/omnuum4106/files/andrewho/files/a_pracitioners_guide_to_growth_models.pdf

Castellano, K. E., & Ho, A. D. (2015). Practical differences among aggregate-level conditional status metrics: From median student growth percentiles to value-added models. *Journal of Educational and Behavioral Statistics*, *40*(1), 35–68. https://doi.org/10.3102/1076998614548485

Chen, X., & Sanchez, E. I. (2024). *Has the predictive validity of high school GPA and ACT scores on postsecondary enrollment changed over time?* (Research Report R2321). ACT. https://www.act.org/content/dam/act/unsecured/documents/pdfs/R2321-Has-the-Predictive-Validity-of-HSGPA-and-ACT-Scores-on-Postsecondary-Enrollment-Changed-2024-01.pdf

Cizek, G. J., & Bunch, M. B. (2006). *Standard setting: A guide to establishing and evaluating performance standards on tests.* Sage. https://doi.org/10.4135/9781412985918

Clinedinst, M. (2015). *2014 state of college admission*. National Association for College Admissions Counseling.

Cohen, J. (1968). Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin, 70*(4), 213–220. https://doi.org/10.1037/h0026256

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297–334. https://doi.org/10.1007/BF02310555

Crouse, J. D., & Allen, J. (2014). College course grades for dual enrollment students. *Community College Journal of Research and Practice*, *38*(6), 494–511. https://doi.org/10.1080/10668926.2011.567168

de Brey, C., Snyder, T. D., Zhang, A., & Dillow, S. A. (2021). *Digest of education statistics 2019* (NCES 2021-009). U.S. Department of Education. https://nces.ed.gov/pubs2021/2021009.pdf

Fleiss, J. L., Levin, B., & Paik, M. C. (2003). *Statistical methods for rates and proportions* (3rd ed.). Wiley & Sons.

Hanson, B. A. (1990). *An investigation of methods for improving estimation of test score distributions* (Research Report 90-4). ACT. https://files.eric.gov/fulltext/ED337464.pdf

Hanson, B. A., & Brennan, R. L. (1990). An investigation of classification consistency indexes estimated under alternative strong true score models. *Journal of Educational Measurement*, *27*(4), 345–359. https://doi.org/10.1111/j.1745-3984.1990.tb00753.x

Harmston, M. (2020). *Increasing your ACT Composite score: Should you test a second time?* (Technical Brief R1813). ACT. https://www.act.org/content/dam/act/unsecured/documents/R1813-should-students-retest-2020-03.pdf

Harmston, M., & Crouse, J. (2016). *Multiple testers: What do we know about them?* [Technical brief]. ACT. https://www.act.org/content/dam/act/unsecured/documents/2016-Tech-Brief-MultipleTesters.pdf

Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities and work environments* (3rd ed.). Psychological Assessment Resources.

Holland, J. L., Whitney, D., Cole, N., & Richards, J. (1969). *An empirical occupational classification derived from a theory of personality and intended for practice and research* (Research Report 29). ACT. https://www.act.org/content/dam/act/unsecured/documents/ACT_RR29.pdf

Houston, M. (1993, April). *Accuracy of validity indices for course placement systems* [Paper presentation]. Annual Meeting of the American Educational Research Association, Atlanta, GA, United States.

Huh, N. R., & Huang, C. Y. (2016). *Examining the validity of ACT Composite scores and high school grade point average for predicting first-year college GPA of special-tested students* (Research Report 2016-7). ACT. https://www.act.org/content/dam/act/unsecured/documents/5931-research-report-2016-7-examining-the-validity-of-act-composite-score-and-hs-gpa.pdf

Joint Committee on Testing Practices. (2004). *Code of fair testing practices in education.* https://www.apa.org/science/programs/testing/fair-testing.pdf

Koenker, R. (2005). *Quantile regression*. Cambridge University Press.

Kolen, M. J. (1984). Effectiveness of analytic smoothing in equipercentile equating. *Journal of Educational Statistics, 9*(1), 25–44. https://doi.org/10.2307/1164830

Kolen, M. J. (1988). Defining score scales in relation to measurement error. *Journal of Educational Measurement, 25*(2), 97–110. https://doi.org/10.1111/j.1745-3984.1988.tb00295.x

Kolen, M. J. (1991). Smoothing methods for estimating test score distributions. *Journal of Educational Measurement, 28*(3), 257–282. https://doi.org/10.1111/j.1745-3984.1991.tb00358.x

Kolen, M. J., & Brennan, R. L. (2014). *Test equating, scaling, and linking: Methods and practices* (3rd ed.). Springer.

Kolen, M. J., & Hanson, B. A. (1989). Scaling the ACT assessment. In R. L. Brennan (Ed.), *Methodology used in scaling the ACT Assessment and PACT+* (pp. 35–55). ACT.

Kolen, M. J., Hanson, B. A., & Brennan, R. L. (1992). Conditional standard errors of measurement for scale scores. *Journal of Educational Measurement, 29*(4), 285–307. https://doi.org/10.1111/j.1745-3984.1992.tb00378.x

Laing, J., & Farmer, M. (1984). *Use of the ACT assessment by examinees with disabilities* (Research Report 84). ACT. https://www.act.org/content/dam/act/unsecured/documents/ACT_RR84.pdf

Li, D., Kapoor, S., Arthur, A., Huang, C.-Y., Cho, Y., Qiu, C., & Wang, H. (2025). *The enhanced ACT linking study report* (Research Report R2515). ACT. https://www.act.org/content/dam/act/unsecured/documents/R2515-The-Enhanced-ACT-Linking-Study-Report-2025-03-updated.pdf

Li, D., Yi, Q., & Harris, D. (2017). *Evidence for paper and online ACT comparability: Spring 2014 and 2015 mode comparability studies* (Research Report 2017-1). ACT. https://files.eric.ed.gov/fulltext/ED583566.pdf

Ling, J., & Radunzel, J. (2017). *Who is likely to graduate high school ready for college?* (Technical Brief R1646). ACT. https://files.eric.ed.gov/fulltext/ED583597.pdf

Livingston, S. A., & Lewis, C. (1995). Estimating the consistency and accuracy of classifications based on test scores. *Journal of Educational Measurement, 32*(2), 179–197. https://doi.org/10.1111/j.1745-3984.1995.tb00462.x

Lorah, J., & Ndum, E. (2013). *Trends in achievement gaps in first-year college courses for racial/ethnic, income, and gender subgroups: A 12-year study* (Research Report 2013-8). ACT. https://www.act.org/content/dam/act/unsecured/documents/ACT_RR2013-8.pdf

Lord, F. M. (1965). A strong true score theory, with applications. *Psychometrika, 30*(3), 239–270. https://doi.org/10.1007/BF02289490

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Addison-Wesley.

Loria, R. N., & Sanchez, E. I. (2024). *Validity of ACT Composite score and high school GPA for predicting probability of timely degree completion: Examining first-year college GPA as a mediator* (Research Report R2419). ACT. https://www.act.org/content/dam/act/unsecured/documents/R2419-Validity-ACT-Composite-Score-HSGPA-2025-02.pdf

Mattern, K., Burrus, J., Camara, W., O'Connor, R., Hansen, M. A., Gambrell, J., Casillas, A., & Bobek, B. (2014). *Broadening the definition of college and career readiness: A holistic approach* (Research Report 2014-5). ACT. https://www.act.org/content/dam/act/unsecured/documents/ACT_RR2014-5.pdf

Mattern, K., Patterson, B. F., Shaw, E. J., Kobrin, J. L., & Barbuti, S. M. (2008). *Differential validity and prediction of the SAT* (Research Report 2008-4). College Board. https://files.eric.ed.gov/fulltext/ED562614.pdf

Mattern, K., Radunzel, J., Bertling, M., & Ho., A. D. (2018). How should colleges treat multiple admissions test scores? *Educational Measurement: Issues and Practice*, *37*(3), 11–23. https://doi.org/10.1111/emip.12199

Mattern, K., Sanchez, E. I., & Ndum, E. (2017). Why do achievement measures underpredict female academic performance? *Educational Measurement: Issues and Practice. 36*(1), 47–57. https://doi.org/10.1111/emip.12138

McNeish, D. M., Radunzel, J., & Sanchez, E. I. (2015). *A multidimensional perspective of college readiness: Relating student and school characteristics to performance on the ACT* (Research Report 2015-6). ACT. https://www.act.org/content/dam/act/unsecured/documents/ACT_RR2015-6.pdf

Millman, J., & Greene, J. (1989). The specification and development of tests of achievement and ability. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 335–366). American Council on Education; Macmillan Publishing Company.

Moore, J. L., Huang, C., Huh, N., Li, T., & Camara, W. (2018). *Testing supports for English learners: A literature review and preliminary ACT research findings* (Working Paper

R1696). ACT. https://www.act.org/content/dam/act/unsecured/documents/R1696-el-supports-2018-05.pdf

Moore, J. L., Li, T., & Lu, Y. (2020). *Reliability of English learners' test scores* (Technical Brief R1754). ACT. https://www.act.org/content/dam/act/unsecured/documents/R1754-el-score-reliability-2020-05.pdf

Moore, J. L., & Schnieders, J. Z. (2025). *Predicting college completion of students who take the ACT with accommodations* (Research Report R2333). ACT. https://www.act.org/content/dam/act/unsecured/documents/R2333-Predicting-College-Completion-of-Students-Who-Take-the-ACT-With-Accommodations-2025-02.pdf

Moore, J. L., & Schnieders, J. Z. (2026a). *Benefits of testing supports for English learners taking the ACT test: Evidence from one state* (Data Byte R2451). ACT. https://www.act.org/content/dam/act/unsecured/documents/r2541-benefits-of-testing-supports-for-english-learners-taking-the-act-test-evidence-from-one-state-2026-04.pdf

Moore, J. L., & Schnieders, J. Z. (2026b). *Coursework patterns of English learners in Grade 11: Evidence from one state* (Data Byte R2542). ACT. https://www.act.org/content/dam/act/unsecured/documents/r2542-coursework-patterns-of-english-learners-in-grade-11-evidence-from-one-state-2026-05.pdf

Moore, J. L., & Schnieders, J. Z. (2026c). *Exploring relationships between English proficiency and ACT test performance of English learners: State 1* (Research Report R2538). ACT. https://www.act.org/content/dam/act/unsecured/documents/r2538-english-proficiency-act-test-performance-els-state-1-2026-04.pdf

Moore, J. L., & Schnieders, J. Z. (2026d). *Relationships between English language proficiency and ACT test performance of English learners: Results from two states* (Data Byte R2450). ACT. https://www.act.org/content/dam/act/unsecured/documents/r2540-relationships-between-english-language-proficiency-and-act-test-performance-of-english-learners-results-from-two-states-2026-05.pdf

Moore, J. L., Schnieders, J. Z., Li, D., & Qiu, C. (2026). *Exploring relationships between English proficiency and ACT test performance of English learners: State 2* (Research Report R2539). ACT. https://www.act.org/content/dam/act/unsecured/documents/R2539-English-proficiency-ACT-test-performance-ELs-State-2-2026-04.pdf

Moore, R., Sanchez, E. I., & San Pedro, S. (2019). *College entrance exams: How does test preparation affect retest scores?* (Research Report R1743). ACT. https://www.act.org/content/dam/act/unsecured/documents/R1743-test-prep-retest-scores-2019-07.pdf

National Institute of Standards and Technology. (2017). Information technology laboratory: Computer security resource center. http://csrc.nist.gov/publications/PubsSPs.html

NCME Ad Hoc Committee on the Development of a Code of Ethics. (2023). *Code of professional responsibilities in educational measurement.* National Council on Measurement in Education. https://ncme.org/wp-content/uploads/2025/10/NCME_STUC_2023_CodeofProfResp_Final_Updated.pdf

Ndum, E. A., Allen, J., Way, J. D., & Casillas, A. (2015). *Effects of psychosocial factors on gender differences in first-year college English and mathematics grades* (Working Paper 2015-9). ACT. https://www.act.org/content/dam/act/unsecured/documents/Working-Paper-2015-09-Effects-of-Psychosocial-Factors-on-Gender-Differences.pdf

Noble, J. P., Crouse, J., Sawyer, R., & Gillespie, M. (1992, April). *Ethnic/gender bias and the differential performance hypothesis: Implications for performance on the ACT assessment* [Paper presentation]. Annual meeting of the American Educational Research Association, San Francisco, CA, United States.

Noble, J. P., Crouse, J., & Schulz, M. (1995, April). *Differential prediction/impact in course placement for ethnic and gender groups* [Paper presentation]. Annual meeting of the National Council on Measurement in Education, San Francisco, CA, United States.

Noble, J. P., Crouse, J., & Schultz, M. (1996). *Differential prediction/impact in course placement for ethnic and gender groups* (Research Report 96-8). ACT. https://www.act.org/content/dam/act/unsecured/documents/ACT_RR96-08.pdf

Noble, J. P., Davenport, M., & Sawyer, R. (2001, April). *Relationships between noncognitive characteristics, high school course work and grades, and performance on a college admissions test* [Paper presentation]. Annual Meeting of the American Education Research Association, Seattle, WA, United States.

Noble, J. P., Davenport, M., Schiel, J., & Pommerich, M. (1999a). *Relationships between the noncognitive characteristics, high school course work and grades, and test scores for*

*ACT-tested students* (Research Report 99-4). ACT. https://www.act.org/content/dam/act/unsecured/documents/ACT_RR99-04.pdf

Noble, J. P., Davenport, M., Schiel, J., & Pommerich, M. (1999b). *High school academic and noncognitive variables related to the ACT scores of racial/ethnic and gender groups* (Research Report 99-6). ACT. https://www.act.org/content/dam/act/unsecured/documents/ACT_RR99-06.pdf

Noble, J. P., & McNabb, T. (1989). *Differential coursework and grades in high school: Implications for performance on the ACT assessment* (Research Report 89-5). ACT. https://files.eric.ed.gov/fulltext/ED322159.pdf

Noble, J. P., Roberts, B., & Sawyer, R. L, (2006). *Student achievement, behavior, perceptions and other factors affecting ACT scores* (Research Report 2006-1). ACT. https://www.act.org/content/dam/act/unsecured/documents/ACT_RR2006-1.pdf

Noble, J. P., & Sawyer, R. (2002). *Predicting different levels of academic success in college using high school GPA and ACT Composite score* (Research Report 2002-4). ACT. https://www.act.org/content/dam/act/unsecured/documents/ACT_RR2002-4.pdf

Noble, J. P., & Sawyer, R. (2013). *A study of the effectiveness of developmental courses for improving success in college* (Research Report 2013-1). ACT. https://www.act.org/content/dam/act/unsecured/documents/ACT_RR2013-1.pdf

Noble, J. P., Schiel, J. L., & Sawyer, R. L. (2004). Assessment and college course placement: Matching students with appropriate instruction. In J. E. Wall & G. R. Walz (Eds.), *Measuring up: Assessment issues for teachers, counselors, and administrators* (pp. 297–311). ERIC Counseling & Student Services Clearinghouse; National Board of Certified Counselors.

Noble, J. P., & Schnelker, D. L. (2007). *Using hierarchical modeling to examine course work and ACT score relationships across high schools* (Research Report 2007-2). ACT. https://www.act.org/content/dam/act/unsecured/documents/ACT_RR2007-2.pdf

Nye, C., Su, R., Rounds, J., & Drasgow, F. (2012). Vocational interests and performance: A quantitative summary of over 60 years of research. *Perspectives on Psychological Science, 7*(4), 384–403. https://doi.org/10.1177/1745691612449021

Patterson, B. F., & Mattern, K. (2012). *Validity of the SAT for predicting first-year grades: 2009 SAT validity sample* (Statistical Report 2012-2). College Board. https://files.eric.ed.gov/fulltext/ED563103.pdf

Payne, J. S., & Allen, J. (2019). *An early look at ACT Academy usage and effectiveness* (Research Report R1785). ACT. https://www.act.org/content/dam/act/unsecured/documents/R1785-academy-first-look-2019-11.pdf

Pilcher, J. K. (1994). The value-driven meaning of grades. *Educational Assessment, 2*(1), 69–88. https://doi.org/10.1207/s15326977ea0201_4

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin*, *135*(2), 322–338. https://doi.org/10.1037/a0014996

Powers, S., Li, D., Suh, H., & Harris, D. J. (2016). *ACT reporting category interpretation guide: Version 1.0*. ACT. https://files.eric.ed.gov/fulltext/ED573778.pdf

Prediger, D. (1996). Alternative dimensions for the Tracey–Rounds interest sphere. *Journal of Vocational Behavior*, *48*(1), 59–67. https://doi.org/10.1006/jvbe.1996.0005

Prediger, D., & Swaney, K. (2004). Work task dimensions underlying the world of work: Research results for diverse occupational databases. *Journal of Career Assessment, 12*(4), 440–459. https://doi.org/10.1177/1069072704267737

Radunzel, J. (2012, October). *Where are 2003 high school graduates seven years later?* [Paper presentation]. Eighth Annual National Symposium on Student Retention, New Orleans, LA, United States. http://www.act.org/content/dam/act/unsecured/documents/WhereDidHighSchoolGraduatesGo.pdf

Radunzel, J. (2019). *Relating ACT writing scores to first-year English composition grades* (Technical Brief R1798). ACT. https://files.eric.ed.gov/fulltext/ED602031.pdf

Radunzel, J., & Allen, J. (2020). *Predicting success on Advanced Placement exams using ACT Aspire, PreACT, and ACT test scores* (Research Report R1835). ACT. https://files.eric.ed.gov/fulltext/ED610228.pdf

Radunzel, J., & Mattern, K. (2020). *Predicting students' chances of completing a degree: How does superscoring compare to other scoring methods when applicants retest?* (Working

Paper R1825). ACT. https://www.act.org/content/dam/act/unsecured/documents/R1825-degree-completion-superscore-2020-05.pdf

Radunzel, J., Mattern, K., Crouse, J., & Westrick, P. (2015). *Development and validation of a STEM Benchmark based on the ACT STEM score* [Technical brief]. ACT. https://www.act.org/content/dam/act/unsecured/documents/2015-Tech-Brief-Development-and-Validation.pdf

Radunzel, J., Mattern, K., & Westrick, P. (2016). *The role of academic preparation and interest on STEM success* (Research Report 2016-8). ACT. https://files.eric.ed.gov/fulltext/ED581664.pdf

Radunzel, J., & Noble, J. (2012). *Predicting long-term college success through degree completion using ACT Composite score, ACT Benchmarks, and high school grade point average* (Research Report 2012-5). ACT. https://www.act.org/content/dam/act/unsecured/documents/ACT_RR2012-5.pdf

Radunzel, J., & Noble, J. (2013). *Differential effects on student demographic groups of using ACT college readiness assessment Composite score, ACT Benchmarks, and high school grade point average for predicting long-term college success through degree completion* (Research Report 2013-5). ACT. https://www.act.org/content/dam/act/unsecured/documents/ACT_RR2013-5.pdf

Radunzel, J., Noble, J., & Wheeler, S. (2014). *Dual-credit/dual-enrollment coursework and long-term college success in Texas* [Issue brief]. ACT. https://www.act.org/content/dam/act/unsecured/documents/DualCreditTexasReport.pdf

Radunzel, J., Westrick, P., Bassiri, D., & Li, D. (2017). *Development and validation of a preliminary ELA readiness benchmark based on the ACT ELA score* (Data Byte R1651). ACT. https://www.act.org/content/dam/act/unsecured/documents/R1651-preliminary-ela-benchmark-2017-06.pdf

Sanchez, E. I. (2013). *Differential effects of using ACT college readiness assessment scores and high school GPA to predict first-year college GPA among racial/ethnic, gender, and income groups* (Research Report 2013-4). ACT. https://www.act.org/content/dam/act/unsecured/documents/ACT_RR2013-4.pdf

Sanchez, E. I. (2018). *ACT Online Prep helps students on the ACT* (Data Byte R1721). ACT. https://www.act.org/content/dam/act/unsecured/documents/R1721-aop-impact-2018-11.pdf

Sanchez, E. I. (2019). *Can using ACT Online Prep improve score gains?* (Technical Brief R1740). ACT. https://www.act.org/content/dam/act/unsecured/documents/R1740-AOP-usage-2019-05.pdf

Sanchez, E. I. (2020). *ACT Online Prep: An effective way to raise ACT scores* (Data Byte R1807). ACT. https://www.act.org/content/dam/act/unsecured/documents/R1807-aop-hours-2020-01.pdf

Sanchez, E. I. (2023). *Evidence of grade inflation since 2010 in high school English, mathematics, social studies, and science courses* (Research Report R2300). ACT. https://www.act.org/content/dam/act/secured/documents/Evidence-of-Grade-Inflation-in-English-Math-Social-Studies-and-Science.pdf

Sanchez, E. I. (2024a). *Has the relationship between college readiness measures and developmental course placement changed in recent years?* (Research Report R2334). ACT. https://files.eric.ed.gov/fulltext/ED673783.pdf

Sanchez, E. I. (2024b). *How has the predictive validity of HSGPA and ACT Composite score on cumulative first-year hours earned changed between 2018 and 2022?* (Research Report R2414). ACT. https://files.eric.ed.gov/fulltext/ED673788.pdf

Sanchez, E. I. (2024c). *Predicting academic success in college: The comparative strength of high school GPA, ACT score, and demographic factors* (Research Report R2421). ACT. https://www.act.org/content/dam/act/unsecured/documents/R2421-Predicting-Academic-Success-in-College-2024-12.pdf

Sanchez, E. I. (2024d). *Predictive validity of high school GPA and ACT Composite score on graduating college GPA: Examining first-year college GPA as a mediator* (Research Report R2417). ACT. https://files.eric.ed.gov/fulltext/ED673778.pdf

Sanchez, E. I. (2025a). *The implications of grade inflation: Evaluating the prediction of first-year college GPA using high school GPA and ACT scores* (Research Report R2432). ACT. https://www.act.org/content/dam/act/unsecured/documents/R2432-The-Implications-of-Grade-Inflation-2025-11.pdf

Sanchez, E. I. (2025b). *The power of preparation: How rigorous high school courses relate to ACT Composite scores* (Research Report R2502). ACT. https://www.act.org/content/dam/act/unsecured/documents/R2502-Power-Preparation-Rigorous-HS-Courses-Enhance-ACTC-Scores-09-2025.pdf

Sanchez, E. I. (2025c). *Predicting ACT ELA and STEM Benchmark attainment: The role of high school coursework and grades* (Research Report R2433). ACT. https://files.eric.ed.gov/fulltext/ED677928.pdf

Sanchez, E. I. (2025d). *Predicting STEM achievement: A comparative study of ACT scores and high school GPA* (Research Report R2501). ACT. https://files.eric.ed.gov/fulltext/ED673777.pdf

Sanchez, E. I., & Buddin, R. (2016). *How accurate are self-reported high school courses, course grades, and grade point average?* (Research Report 2016-3). ACT. https://www.act.org/content/dam/act/unsecured/documents/5269-research-report-how-accurate-are-self-reported-hs-courses.pdf

Sanchez, E. I., & Moore, R. (2022). *Grade inflation continues to grow in the past decade* (Research Report R2134). ACT. https://www.act.org/content/dam/act/secured/documents/pdfs/Grade-Inflation-Continues-to-Grow-in-the-Past-Decade-Final-Accessible.pdf

Sawyer, R. (1985). *Using demographic information in predicting college freshman grades* (Research Report 87). ACT. https://www.act.org/content/dam/act/unsecured/documents/ACT_RR87.pdf

Sawyer, R. (1989). Research design and sampling for the Academic Skills Study. In R. L. Brennan (Ed.), *Methodology used in scaling the ACT assessment and PACT+* (pp. 19–33). ACT.

Sawyer, R. (2008). *Benefits of additional high school course work and improved course performance in preparing students for college* (Research Report 2008-1). ACT. https://www.act.org/content/dam/act/unsecured/documents/ACT_RR2008-1.pdf

Sawyer, R. (2010). *Usefulness of high school average and ACT scores in making college admission decisions* (Research Report 2010-2). ACT. https://www.act.org/content/dam/act/unsecured/documents/ACT_RR2010-2.pdf

Sawyer, R. (2013a). Beyond correlations: Usefulness of high school GPA and test scores in making college admissions decisions. *Applied Measurement in Education*, *26*(2), 89–112. https://doi.org/10.1080/08957347.2013.765433

Sawyer, R. (2013b). *Interpreting changes over time in high school average, ACT college readiness assessment Composite scores, and ACT College Readiness Benchmark attainment rates* (Research Report 2013-9). ACT. https://www.act.org/content/dam/act/unsecured/documents/ACT_RR2013-9.pdf

Sawyer, R., Laing, J., & Houston, M. (1988). *Accuracy of self-reported high school courses and grades of college-bound students* (Research Report 88-1). ACT. https://www.act.org/content/dam/act/unsecured/documents/ACT_RR88-1.pdf

Schiel, J. (1998). *Academic benefits in high school of an intensive summer program for academically talented seventh graders* (Research Report 98-4). ACT. https://files.eric.ed.gov/fulltext/ED424262.pdf

Schiel, J., & Noble, J. (1992). *The effects of data truncation on estimated validity indices for course placement* (Research Report 92-3). ACT. https://www.act.org/content/dam/act/unsecured/documents/ACT_RR92-03.pdf

Schiel, J., Pommerich, M., & Noble, J. P. (1996). *Factors associated with longitudinal educational achievement, as measured by PLAN and ACT assessment scores* (Research Report 96-5). ACT. https://files.eric.ed.gov/fulltext/ED403298.pdf

Schiel, J. L. (2020). *Test-related behaviors and performance on the ACT* (Issue Brief R1742). ACT. https://files.eric.ed.gov/fulltext/ED604111.pdf

Schiel, J. L., & Harmston, M. (2000). *Validating two-stage course placement systems when data are truncated* (Research Report 2000-3). ACT. https://www.act.org/content/dam/act/unsecured/documents/ACT_RR2000-3.pdf

Schiel, J. L., & Valiga, M. J. (2014a). *How ACT scores change with test preparation* (Information Brief 2014-6). ACT. https://www.act.org/content/dam/act/unsecured/documents/Info-Brief-2014-6.pdf

Schiel, J. L., & Valiga, M. J. (2014b). *How the amount of time spent on test preparation relates to ACT scores* (Information Brief 2014-7). ACT. https://www.act.org/content/dam/act/unsecured/documents/Info-Brief-2014-7-rev1.pdf

Schiel, J. L., & Valiga, M. J. (2014c). *How particular outcomes of test preparation relate to ACT scores* (Information Brief 2014-8). ACT. https://www.act.org/content/dam/act/unsecured/documents/Info-Brief-2014-8.pdf

Scholes, R. J., & Lain, M. M. (1997, March). *The effects of test preparation activities on ACT assessment scores* [Paper presentation]. Annual Meeting of the American Educational Research Association, Chicago, IL, United States.

Scholes, R. J., & McCoy, T. R. (1998, April). *The effects of type, length, and content of test preparation activities on ACT assessment scores* [Paper presentation]. Annual Meeting of the American Educational Research Association, San Diego, CA, United States.

Snyder, T. D., & Dillow, S. A. (2013). *Digest of education statistics 2012* (NCES 2014-015). U.S. Department of Education. https://nces.ed.gov/pubs2014/2014015.pdf

Sparks, D., & Malkus, N. (2013). *First-year undergraduate remedial coursetaking: 1999-2000, 2003-04, 2007-08* (Statistics in Brief, NCES 2013-013). National Center for Education Statistics. https://www.air.org/sites/default/files/2021-06/First_Year_Undergraduate_Remedial_Counseling_Jan2013_0.pdf

Steedle, J., Pashley, P., & Cho, YoungWoo. (2020). *Three studies of comparability between paper-based and computer-based testing for the ACT* (Research Report R1847). ACT. https://www.act.org/content/dam/act/unsecured/documents/R1847-three-comparability-studies-2020-12.pdf

Stiggins, R. J., Frisbie, D. A., & Griswold, P. A. (1989). Inside high school grading practices: Building a research agenda. *Educational Measurement: Issues and Practice, 8*(2), 5–14. https://doi.org/10.1111/j.1745-3992.1989.tb00315.x

Stocking, M. L., & Lord, F. M. (1983). Developing a common metric in item response theory. *Applied Psychological Measurement*, *7*(2), 201–210. https://doi.org/10.1177/014662168300700208

Tracey, T., & Robbins, S. (2006). The interest-major congruence and college success relation: A longitudinal study. *Journal of Vocational Behavior*, *69*(1), 64–89. https://doi.org/10.1016/j.jvb.2005.11.003

United States Department of Education. (1994). *What do student grades mean? Differences across schools* (Research Report OR-94-3401). Office of Educational Research and Improvement. https://files.eric.ed.gov/fulltext/ED367666.pdf

Valiga, M. (1986). *Accuracy of self-reported high school course and grade information* (Research Report 87-1). ACT. https://www.act.org/content/dam/act/unsecured/documents/ACT-RR87-01.pdf

Wai, J. & Allen, J. (2019). What boosts talent development? Examining predictors of academic growth in secondary school among academically advanced youth across 21 years. *Gifted Child Quarterly, 63*(4), 253–272. https://doi.org/10.1177/0016986219869042

Wang, L., & Steedle, J. (2020). *An investigation of differential mode effects when comparing paper and online ACT testing* (Technical Brief 1838). ACT. https://files.eric.ed.gov/fulltext/ED610121.pdf

Wang, M. W., & Stanley, J. L. (1970). Differential weighting: A review of methods and empirical studies. *Review of Educational Research*, *40*(5), 663–705. https://doi.org/10.3102/00346543040005663

Webb, N. L. (2002, March 28). *Depth-of-knowledge levels for four content areas* [Unpublished manuscript]. http://ossucurr.pbworks.com/w/file/fetch/49691156/norm%20web%20dok%20by%20subject%20area.pdf

Westrick, P. A. (2016). *The joint use of ACT scores and high school grade point average for predicting success at community colleges* [Technical brief]. ACT. https://www.act.org/content/dam/act/unsecured/documents/5582-tech-brief-joint-use-of-act-scores-and-hs-grade-point-average.pdf

Westrick, P. A. (2017). Reliability estimates for undergraduate grade point average. *Educational Assessment*, *22*(4), 231–252. https://doi.org/10.1080/10627197.2017.1381554

Westrick, P. A., & Allen, J. (2014). *Validity evidence for ACT Compass placement tests* (Research Report 2014-2). ACT. https://files.eric.ed.gov/fulltext/ED546849.pdf

Westrick, P. A., Le, H., Robbins, S. B., Radunzel, J., & Schmidt, F. L. (2015). College performance and retention: A meta-analysis of the predictive validities of ACT scores, high school grades, and SES. *Educational Assessment*, *20*(1), 23–45. https://doi.org/10.1080/10627197.2015.997614

Westrick, P. A., Schmidt, F., Le, H., Robbins, S., & Radunzel, J. (2021). The road to retention passes through first year academic performance: A meta-analytic path analysis of academic performance and persistence. *Educational Assessment*, *26*(1), 35–51. https://doi.org/10.1080/10627197.2020.1848423

Whitney, D. R. (1989). Educational admissions and placement. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 515–525). American Council on Education; Macmillan.

Willingham, W. W., Pollack, J. M., & Lewis, C. (2002). Grades and test scores: Accounting for observed differences. *Journal of Educational Measurement*, *39*(1), 1–37. https://doi.org/10.1111/j.1745-3984.2002.tb01133.x

Woodruff, D. J., & Ziomek, R. L. (2004a). *High school grade inflation from 1991 to 2003* (Research Report 2004-4). ACT. https://www.act.org/content/dam/act/unsecured/documents/ACT_RR2004-4.pdf

Woodruff, D. J., & Ziomek, R. L. (2004b). *Differential grading standards among high schools* (Research Report 2004-2). ACT. https://www.act.org/content/dam/act/unsecured/documents/ACT_RR2004-2.pdf

Zhang, Q., & Sanchez, E. I. (2013). *High school grade inflation from 2004 to 2011* (Research Report 2013-3). ACT. https://www.act.org/content/dam/act/unsecured/documents/ACT_RR2013-3.pdf

Zimowski, M., Muraki, E., Mislevy, R. J., & Bock, R. D. (2003). *BILOG-MG* (Version 3.0) [Computer software+e]. Scientific Software International.

Ziomek, R. L., & Andrews, K. M. (1996). *Predicting the college grade point averages of special-tested students from their ACT assessment scores and high school grades* (Research Report 96-7). ACT. https://www.act.org/content/dam/act/unsecured/documents/ACT_RR96-07.pdf

Ziomek, R. L., & Svec, J. C. (1995). *High school grades and achievement: Evidence of grade inflation* (Research Report 95-3). ACT. https://www.act.org/content/dam/act/unsecured/documents/ACT_RR95-03.pdf

ACT®